

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

**METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS
AND METHODS OF SCREENING FOR ANGIOGENESIS
MODULATORS**

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority to USSN 09/784,356, filed February 14 2001;
10 USSN 09/791,390, filed February 22, 2001; USSN 60/285,475, filed April 19, 2001, USSN
60/310,025, filed August 3, 2001, and USSN 60/334,244, filed November 29, 2001, each of
which is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

15 The invention relates to the identification of nucleic acid and protein
expression profiles and nucleic acids, products, and antibodies thereto that are involved in
angiogenesis; and to the use of such expression profiles and compositions in diagnosis and
therapy of angiogenesis. The invention further relates to methods for identifying and using
agents and/or targets that modulate angiogenesis.

20

BACKGROUND OF THE INVENTION

Both vasculogenesis, the development of an interactive vascular system
comprising arteries and veins, and angiogenesis, the generation of new blood vessels, play a
role in embryonic development. In contrast, angiogenesis is limited in a normal adult to the
25 placenta, ovary, endometrium and sites of wound healing. However, angiogenesis, or its
absence, plays an important role in the maintenance of a variety of pathological states. Some
of these states are characterized by neovascularization, *e.g.*, cancer, diabetic retinopathy,
glaucoma, and age related macular degeneration. Others, *e.g.*, stroke, infertility, heart
disease, ulcers, and scleroderma, are diseases of angiogenic insufficiency.

30 Angiogenesis has a number of stages (see, *e.g.*, Folkman, *J.Natl Cancer Inst.*
82:4-6, 1990; Firestein, *J Clin Invest.* 103:3-4, 1999; Koch, *Arthritis Rheum.* 41:951-62, 1998;
Carter, *Oncologist* 5(Suppl 1):51-4, 2000; Browder *et al.*, *Cancer Res.* 60:1878-86, 2000; and
Zhu and Witte, *Invest New Drugs* 17:195-212, 1999). The early stages of angiogenesis

include endothelial cell protease production, migration of cells, and proliferation. The early stages also appear to require some growth factors, with VEGF, TGF- α , angiostatin, and selected chemokines all putatively playing a role. Later stages of angiogenesis include population of the vessels with mural cells (pericytes or smooth muscle cells), basement
5 membrane production, and the induction of vessel bed specializations. The final stages of vessel formation include what is known as "remodeling", wherein a forming vasculature becomes a stable, mature vessel bed. Thus, the process is highly dynamic, often requiring coordinated spatial and temporal waves of gene expression.

Conversely, the complex process may be subject to disruption by interfering
10 with one or more critical steps. Thus, the lack of understanding of the dynamics of angiogenesis prevents therapeutic intervention in serious diseases such as those indicated. It is an object of the invention to provide methods that can be used to screen compounds for the ability to modulate angiogenesis. Additionally, it is an object to provide molecular targets for therapeutic intervention in disease states which either have an undesirable excess or a deficit
15 in angiogenesis. The present invention provides solutions to both.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for detecting or modulating angiogenesis associated sequences.

20 In one aspect, the invention provides a method of detecting an angiogenesis-associated transcript in a cell in a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridized to a sequence at least 80% identical to a sequence as shown in Tables 1-8. In one embodiment, the biological sample is a tissue sample. In another embodiment, the biological sample comprises isolated
25 nucleic acids, which are often mRNA.

In another embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide. Often, the polynucleotide comprises a sequence as shown in Tables 1-8. The polynucleotide can be labeled, for example, with a fluorescent label and can be immobilized on a solid
30 surface.

In other embodiments the patient is undergoing a therapeutic regimen to treat a disease associated with angiogenesis or the patient is suspected of having an angiogenesis-associated disorder.

In another aspect, the invention comprises an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1-8. The nucleic acid molecule can be labeled, for example, with a fluorescent label,

5 In other aspects, the invention provides an expression vector comprising an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1-8 or a host cell comprising the expression vector.

In another embodiment, the isolated nucleic acid molecule encodes a polypeptide having an amino acid sequence as shown in Table 8.

10 In another aspect, the invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1-8. In one embodiment, the isolated polypeptide has an amino acid sequence as shown in Table 8.

15 In another embodiment, the invention provides an antibody that specifically binds a polypeptide that has an amino acid sequence as shown in Table 8 or which is encoded by a nucleotide sequence of Tables 1-8. The antibody can be conjugated or fused to an effector component such as a fluorescent label, a toxin, or a radioisotope. In some embodiments, the antibody is an antibody fragment or a humanized antibody.

20 In another aspect, the invention provides a method of detecting a cell undergoing angiogenesis in a biological sample from a patient, the method comprising contacting the biological sample with an antibody that specifically binds to a polypeptide that has an amino acid sequence as shown in Table 8 or which is encoded by a nucleotide sequence of Tables 1-8. In some embodiments, the antibody is further conjugated or fused to an effector component, for example, a fluorescent label.

25 In another embodiment, the invention provides a method of detecting antibodies specific to angiogenesis in a patient, the method comprising contacting a biological sample from the patient with a polypeptide which is encoded by a nucleotide sequence of Tables 1-8.

30 The invention also provides a method of identifying a compound that modulates the activity of an angiogenesis-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a polypeptide that comprises at least 80% identity to an amino acid sequence as shown in Table 8 or which is encoded by a nucleotide sequence of Tables 1-8; and (ii) detecting an increase or a decrease in the activity of the polypeptide. In one embodiment, the polypeptide has an amino acid sequence as shown in Table 8 or is a

polypeptide encoded by a nucleotide sequence of Tables 1-8. In another embodiment, the polypeptide is expressed in a cell.

The invention also provides a method of identifying a compound that modulates angiogenesis, the method comprising steps of: (i) contacting the compound with a cell undergoing angiogenesis; and (ii) detecting an increase or a decrease in the expression of a polypeptide sequence as shown in Table 8 or a polypeptide which is encoded by a nucleotide sequence of Tables 1-8. In one embodiment, the detecting step comprises hybridizing a nucleic acid sample from the cell with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-8. In another embodiment, the method further comprises detecting an increase or decrease in the expression of a second sequence as shown in Table 8 or a polypeptide which is encoded by a nucleotide sequence of Tables 1-8.

In another embodiment, the invention provides a method of inhibiting angiogenesis in a cell that expresses a polypeptide at least 80% identical to a sequence as shown in Table 8 or which is 80% identical to a polypeptide encoded by a nucleotide sequence of Tables 1-8, the method comprising the step of contacting the cell with a therapeutically effective amount of an inhibitor of the polypeptide. In one embodiment, the polypeptide has an amino acid sequence shown in Table 8 or is a polypeptide which is encoded by a nucleotide sequence of Tables 1-8. In another embodiment, the inhibitor is an antibody.

In other embodiments, the invention provides a method of activating angiogenesis in a cell that expresses a polypeptide at least 80% identical to a sequence as shown in Table 8 or at least 80% identical to a polypeptide which is encoded by a nucleotide sequence of Tables 1-8, the method comprising the step of contacting the cell with a therapeutically effective amount of an activator of the polypeptide. In one embodiment, the polypeptide has an amino acid sequence shown in Table 8 or is a polypeptide which is encoded by a nucleotide sequence of Tables 1-8.

Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

Tables 1-8 provide nucleotide sequence of genes that exhibit changes in expression levels as a function of time in tissue undergoing angiogenesis compared to tissue that is not.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and treatment of disorders associated with angiogenesis (sometimes referred to herein as angiogenesis disorders or AD), as well as methods for screening for compositions which modulate angiogenesis. By "disorder associated with angiogenesis" or "disease associated with angiogenesis" herein is meant a disease state which is marked by either an excess or a deficit of blood vessel development. Angiogenesis disorders associated with increased angiogenesis include, but are not limited to, cancer and proliferative diabetic retinopathy. Pathological states for which it may be desirable to increase angiogenesis include stroke, heart disease, infertility, ulcers, wound healing, ischemia, and sclerodoma. Solid tumors typically require angiogenesis to support or sustain growth, e.g., breast, colon, lung, brain, bladder, and prostate tumors. Other AD include, e.g., arthritis, inflammatory bowel disease, diabetes retinopathy, macular degeneration, atherosclerosis, and psoriasis. Also provided are methods for treating AD.

15 Definitions

The term "angiogenesis protein" or "angiogenesis polynucleotide" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acids, to an angiogenesis protein sequence of Table 8; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence of Table 8, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to an anti-sense strand corresponding to a nucleic acid sequence of Tables 1-8 and conservatively modified variants thereof; (4) have a nucleic acid sequence that has greater than about 95%, preferably greater than about 96%, 97%, 98%, 99%, or higher nucleotide sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a sense sequence corresponding to one set out in Tables 1-8. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or any mammal. An "angiogenesis polypeptide" and an "angiogenesis polynucleotide," include both naturally occurring or recombinant.

A "full length" angiogenesis protein or nucleic acid refers to an angiogenesis polypeptide or polynucleotide sequence, or a variant thereof, that contains all of the elements normally contained in one or more naturally occurring, wild type angiogenesis polynucleotide or polypeptide sequences. The "full length" may be prior to, or after, various stages of post-translation processing.

"Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, *e.g.*, of an angiogenic protein. Such samples include, but are not limited to, tissue isolated from primates, *e.g.*, humans, or rodents, *e.g.*, mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate *e.g.*, chimpanzee or human; cow; dog; cat; a rodent, *e.g.*, guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (*e.g.*, isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention *in vivo*. Archival tissues, having treatment or outcome history, will be particularly useful.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, about 70% identity, preferably 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region (*e.g.*, SEQ ID NOS:1-229), when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default
5 program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20
10 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and
20 visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

A preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1997) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectively. BLAST and BLAST 2.0 are used, with the
25 parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying
30 short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as

far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be

prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture Collection catalog or web site, www.atcc.org).

The terms "polypeptide," "peptide" and "protein" are used interchangeably
5 herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids,
10 as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is
15 bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an
20 amino acid, but that functions in a manner similar to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

25 "Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large
30 number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified

variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g., Creighton, Proteins* (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g., Alberts et al., Molecular Biology of the Cell* (3rd ed., 1994) and Cantor and Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional

structure formed, usually by the noncovalent association of independent tertiary units.

Anisotropic terms are also known as energy terms.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ^{32}P , fluorescent dyes, electron-dense reagents, enzymes (*e.g.*, as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins which can be made detectable, *e.g.*, by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody. The "effector" can be a variety of molecules including, for example, detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" *e.g.*, beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, *e.g.*, biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.*, A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin

complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

The term "recombinant" when used with reference, *e.g.*, to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been
5 modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

10 The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, *e.g.*, a promoter from one source and a coding region
15 from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (*e.g.*, a fusion protein).

A "promoter" is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type
20 promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is
25 active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

30 An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

5 The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in
10 Tijssen, *Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50%
15 of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to
20 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS,
25 incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C. For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C
30 to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification

reactions are provided, *e.g.*, in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical.

- 5 This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice
10 background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, *e.g.*, and Current Protocols in Molecular Biology, ed. Ausubel, *et al*

- The phrase "functional effects" in the context of assays for testing compounds
15 that modulate activity of an angiogenesis protein includes the determination of a parameter that is indirectly or directly under the influence of the angiogenesis protein, *e.g.*, a functional, physical, or chemical effect, such as the ability to increase or decrease angiogenesis. It includes binding activity, the ability of cells to proliferate, expression in cells undergoing angiogenesis, and other characteristics of angiogenic cells. "Functional effects" include *in vitro*, *in vivo*, and *ex vivo* activities.
20

- By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of an angiogenesis protein sequence, *e.g.*, functional, physical and chemical effects. Such functional effects can be measured by any means known to those skilled in the art, *e.g.*,
25 changes in spectroscopic characteristics (*e.g.*, fluorescence, absorbance, refractive index), hydrodynamic (*e.g.*, shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the angiogenesis protein; measuring binding activity or binding assays, *e.g.* binding to antibodies, and measuring cellular proliferation, particularly endothelial cell proliferation, cell viability, cell division
30 especially of endothelial cells, lumen formation and capillary or vessel growth or formation. Determination of the functional effect of a compound on angiogenesis can also be performed using angiogenesis assays known to those of skill in the art such as an *in vitro* assays, *e.g.*, *in vitro* endothelial cell tube formation assays, and other assays such as the chick CAM assay, the mouse corneal assay, and assays that assess vascularization of an implanted tumor. The

functional effects can be evaluated by many means known to those skilled in the art, *e.g.*, microscopy for quantitative or qualitative measures of alterations in morphological features, *e.g.*, tube or blood vessel formation, measurement of changes in RNA or protein levels for angiogenesis-associated sequences, measurement of RNA stability, identification of
5 downstream or reporter gene expression (CAT, luciferase, β -gal, GFP and the like), *e.g.*, via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

"Inhibitors", "activators", and "modulators" of angiogenic polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules
10 identified using *in vitro* and *in vivo* assays of angiogenic polynucleotide and polypeptide sequences. Inhibitors are compounds that, *e.g.*, bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of angiogenesis proteins, *e.g.*, antagonists. "Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate
15 angiogenesis protein activity. Inhibitors, activators, or modulators also include genetically modified versions of angiogenesis proteins, *e.g.*, versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, *e.g.*, expressing the angiogenic protein *in vitro*, in cells, or cell membranes, applying putative modulator
20 compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of angiogenesis can also be identified by incubating angiogenic cells with the test compound and determining increases or decreases in the expression of 1 or more angiogenesis proteins, *e.g.*, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more angiogenesis proteins, such as angiogenesis proteins comprising the sequences set out in Table 8.

25 Samples or assays comprising angiogenesis proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%,
30 preferably 50%, more preferably 25-0%. Activation of an angiogenesis polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (*i.e.*, two to five fold higher relative to the control), more preferably 1000-3000% higher.

“Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody will be most critical in specificity and affinity of binding.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see *Fundamental Immunology* (Paul ed., 3d ed. 1993)). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty *et al.*, *Nature* 348:552-554 (1990)).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)).

Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (*see, e.g., McCafferty et al., Nature* 348:552-554 (1990); Marks *et al., Biotechnology* 10:779-783 (1992)).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.,* an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

The detailed description of the invention includes discussion of the following aspects of the invention:

- Expression of angiogenesis-associated sequences
- Informatics
- Angiogenesis-associated sequences
- Detection of angiogenesis sequence for diagnostic and therapeutic applications
- Modulators of angiogenesis
- Methods of identifying variant angiogenesis-associated sequences
- Administration of pharmaceutical and vaccine compositions
- Kits for use in diagnostic and/or prognostic applications.

Expression of angiogenesis-associated sequences

In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is unique to the state of the cell. That is, normal tissue may be distinguished from AD tissue. By comparing expression profiles of tissue in known different angiogenesis states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. The identification of sequences that are

differentially expressed in angiogenic versus non-angiogenic tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-regulate angiogenesis, and thus tumor growth or recurrence, in a particular patient. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Angiogenic tissue can also be analyzed to determine the stage of angiogenesis in the tissue. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; for example, screening can be done for drugs that suppress the angiogenic expression profile. This may be done by making biochips comprising sets of the important angiogenesis genes, which can then be used in these screens. These methods can also be done on the protein basis; that is, protein expression levels of the angiogenic proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the angiogenic nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the angiogenic proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in angiogenesis, herein termed "angiogenesis sequences". As outlined below, angiogenesis sequences include those that are up-regulated (i.e. expressed at a higher level) in disorders associated with angiogenesis, as well as those that are down-regulated (i.e. expressed at a lower level). In a preferred embodiment, the angiogenesis sequences are from humans; however, as will be appreciated by those in the art, angiogenesis sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other angiogenesis sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.). Angiogenesis sequences from other organisms may be obtained using the techniques outlined below.

Angiogenesis sequences can include both nucleic acid and amino acid sequences. In a preferred embodiment, the angiogenesis sequences are recombinant nucleic acids. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid *e.g.*, using polymerases and endonucleases, in a form not normally found in nature. Thus an isolated nucleic acid, in a linear form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is

understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, *i.e.* using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered
5 recombinant for the purposes of the invention.

Similarly, a "recombinant protein" is a protein made using recombinant techniques, *i.e.* through the expression of a recombinant nucleic acid as depicted above. A recombinant protein is distinguished from naturally occurring protein by at least one or more characteristics. For example, the protein may be isolated or purified away from some or all
10 of the proteins and compounds with which it is normally associated in its wild type host, and thus may be substantially pure. For example, an isolated protein is unaccompanied by at least some of the material with which it is normally associated in its natural state, preferably constituting at least about 0.5%, more preferably at least about 5% by weight of the total protein in a given sample. A substantially pure protein comprises at least about 75% by
15 weight of the total protein, with at least about 80% being preferred, and at least about 90% being particularly preferred. The definition includes the production of an angiogenesis protein from one organism in a different organism or host cell. Alternatively, the protein may be made at a significantly higher concentration than is normally seen, through the use of an inducible promoter or high expression promoter, such that the protein is made at increased
20 concentration levels. Alternatively, the protein may be in a form not normally found in nature, as in the addition of an epitope tag or amino acid substitutions, insertions and deletions, as discussed below.

In a preferred embodiment, the angiogenesis sequences are nucleic acids. As will be appreciated by those in the art and is more fully outlined below, angiogenesis
25 sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; for example, biochips comprising nucleic acid probes to the angiogenesis sequences can be generated. In the broadest sense, then, by "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the
30 present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have alternate backbones, comprising, for example, phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press); and peptide nucleic acid backbones and linkages. Other

analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or
5 more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, for example to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip.

As will be appreciated by those in the art, nucleic acid analogs may find use in
10 the present invention. In addition, mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in
15 contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (T_m) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4°C drop in T_m for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9°C. Similarly,
20 due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated
25 by those in the art, the depiction of a single strand also defines the sequence of the complementary strand; thus the sequences described herein also provide the complement of the sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine,
30 xanthine hypoxanthine, isocytosine, isoguanine, etc. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus for example the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

An angiogenesis sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the angiogenesis sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization
5 conditions.

For identifying angiogenesis-associated sequences, the angiogenesis screen typically includes comparing genes identified in a modification of an *in vitro* model of angiogenesis as described in Hiraoka, Cell 95:365 (1998) with genes identified in controls. Samples of normal tissue and tissue undergoing angiogenesis are applied to biochips
10 comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, for example from Affymetrix. Gene expression profiles as described herein are generated and the data analyzed.

In a preferred embodiment, the genes showing changes in expression as
15 between normal and disease states are compared to genes expressed in other normal tissues, including, but not limited to lung, heart, brain, liver, breast, kidney, muscle, prostate, small intestine, large intestine, spleen, bone and placenta. In a preferred embodiment, those genes identified during the angiogenesis screen that are expressed in any significant amount in other tissues are removed from the profile, although in some embodiments, this is not necessary.
20 That is, when screening for drugs, it is usually preferable that the target be disease specific, to minimize possible side effects.

In a preferred embodiment, angiogenesis sequences are those that are up-regulated in angiogenesis disorders; that is, the expression of these genes is higher in the disease tissue as compared to normal tissue. "Up-regulation" as used herein means at least
25 about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. All accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, *e.g.*, Benson, DA, et al., Nucleic Acids Research 26:1-7 (1998) and <http://www.ncbi.nlm.nih.gov/>. Sequences are also
30 available in other databases, *e.g.*, European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ). In addition, most preferred genes were found to be expressed in a limited amount or not at all in heart, brain, lung, liver, breast, kidney, prostate, small intestine and spleen.

In another preferred embodiment, angiogenesis sequences are those that are down-regulated in the angiogenesis disorder; that is, the expression of these genes is lower in angiogenic tissue as compared to normal tissue. "Down-regulation" as used herein means at least about a two-fold change, preferably at least about a three fold change, with at least about
5 five-fold or higher being preferred.

Angiogenesis sequences according to the invention may be classified into discrete clusters of sequences based on common expression profiles of the sequences. Expression levels of angiogenesis sequences may increase or decrease as a function of time in a manner that correlates with the induction of angiogenesis. Alternatively, expression levels
10 of angiogenesis sequences may both increase and decrease as a function of time. For example, expression levels of some angiogenesis sequences are temporarily induced or diminished during the switch to the angiogenesis phenotype, followed by a return to baseline expression levels. Tables 1-8 provides genes, the mRNA expression of which varies as a function of time in angiogenesis tissue when compared to normal tissue.

15 In a particularly preferred embodiment, angiogenesis sequences are those that are induced for a period of time, typically by positive angiogenic factors, followed by a return to the baseline levels. Sequences that are temporarily induced provide a means to target angiogenesis tissue, for example neovascularized tumors, at a particular stage of angiogenesis, while avoiding rapidly growing tissue that require perpetual vascularization.
20 Such positive angiogenic factors include α FGF, β FGF, VEGF, angiogenin and the like.

Induced angiogenesis sequences also are further categorized with respect to the timing of induction. For example, some angiogenesis genes may be induced at an early time period, such as within 10 minutes of the induction of angiogenesis. Others may be induced later, such as between 5 and 60 minutes, while yet others may be induced for a time
25 period of about two hours or more followed by a return to baseline expression levels.

In another preferred embodiment are angiogenesis sequences that are inhibited or reduced as a function of time followed by a return to "normal" expression levels. Inhibitors of angiogenesis are examples of molecules that have this expression profile. These sequences also can be further divided into groups depending on the timing of diminished
30 expression. For example, some molecules may display reduced expression within 10 minutes of the induction of angiogenesis. Others may be diminished later, such as between 5 and 60 minutes, while others may be diminished for a time period of about two hours or more

followed by a return to baseline. Examples of such negative angiogenic factors include thrombospondin and endostatin to name a few.

In yet another preferred embodiment are angiogenesis sequences that are induced for prolonged periods. These sequences are typically associated with induction of angiogenesis and may participate in induction and/or maintenance of the angiogenesis phenotype.

In another preferred embodiment are angiogenesis sequences, the expression of which is reduced or diminished for prolonged periods in angiogenic tissue. These sequences are typically angiogenesis inhibitors and their diminution is correlated with an increase in angiogenesis.

Informatics

The ability to identify genes that undergo changes in expression with time during angiogenesis can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with angiogenesis-associated disease. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (*see*, Anderson, L., "Pharmaceutical Proteomics: Targets, Mechanism, and Function," paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (*see*, U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (*e.g.*, nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of data assay data. The data contained in the database is acquired, *e.g.*, using array analysis either singly or in a library format. The database can be in substantially any form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for any assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative
5 and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample undergoing angiogenesis, *i.e.*, the identification of angiogenesis-associated sequences described herein, provide an abundance of information, which can be correlated with pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological
10 status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, prior data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational
15 database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects
20 for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences
25 in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S.
30 Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, *e.g.*, with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

5 In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, *e.g.*, a neoplastic lesion or another tissue specimen to be analyzed for angiogenesis. In another variation, the assay records cross-tabulate one or more of the following parameters for each
10 target species in a sample: (1) a unique identification code, which can include, *e.g.*, a target molecular structure and/or characteristic separation coordinate (*e.g.*, electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of
15 target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or
20 transistor gate states, such as an array of cells in a DRAM device (*e.g.*, each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment, the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

25 When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program
30 embodiment thereof (*e.g.*, FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format

(*e.g.*, Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, *etc.*) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

5 The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (*e.g.*, computer, disk array, *etc.*) comprises a pattern of magnetic domains (*e.g.*, magnetic disk) and/or charge domains (*e.g.*, an array of DRAM
10 cells) composing a bit pattern encoding data acquired from an assay of the invention.

 The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a
15 database comprising a plurality of assay results obtained by the method of the invention.

 In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably
20 initialized to load and execute the computer program for alignment and/or comparison of the assay results. Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

 The target data or record and the computer program can be transferred to
25 secondary memory, which is typically random access memory (*e.g.*, DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (*e.g.*, binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (*e.g.*, Intel Pentium, PowerPC, Alpha,
30 PA-8000, SPARC, MIPS 4400, MIPS 10000, VAX, *etc.*); a program can be a commercial or public domain molecular biology software package (*e.g.*, UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (*e.g.*, DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, *etc.*); an I/O device can be a terminal comprising a video display and a keyboard, a modem,

an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

10 *Angiogenesis-associated sequences*

Angiogenesis proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the angiogenesis protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus or associated with the intracellular side of the plasma membrane. Intracellular proteins are involved in all aspects of cellular function and replication (including, *e.g.*, signaling pathways); aberrant expression of such proteins often results in unregulated or disregulated cellular processes (see, *e.g.*, Molecular Biology of the Cell, 3rd Edition, Alberts, Ed., Garland Pub., 1994). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of primary

sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate.

In another embodiment, the angiogenesis sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell.

- 5 They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine
- 10 kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

- Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor
- 15 guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of
- 20 transmembrane domains include approximately 20 consecutive hydrophobic amino acids that may be followed or flanked by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein may be predicted (see, *e.g.* PSORT web site <http://psort.nibb.ac.jp/>).

- The extracellular domains of transmembrane proteins are diverse; however,
- 25 conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like.
- 30 For example, growth factors such as EGF, FGF and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell for example via a glycosylphosphatidylinositol

(GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Angiogenesis proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for immunotherapeutics, as are
5 described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ*. Alternatively, antibodies can also label intracellular proteins, in which case samples are typically permeabilized to provide access to intracellular proteins.

It will also be appreciated by those in the art that a transmembrane protein can
10 be made soluble by removing transmembrane sequences, for example through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the angiogenesis proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins have a signal
15 peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; by virtue of their circulating nature, they serve to transmit signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor) or an endocrine manner (acting
20 on cells at a distance). Thus secreted molecules find use in modulating or altering numerous aspects of physiology. Angiogenesis proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, *e.g.*, for blood or serum tests.

An angiogenesis sequence is typically initially identified by substantial nucleic
25 acid and/or amino acid sequence homology or linkage to the angiogenesis sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

30 As detailed in the definitions, percent identity can be determined using an algorithm such as BLAST. A preferred method utilizes the BLASTN module of WU-BLAST-2 set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively. The alignment may include the introduction of gaps in the sequences to be aligned. In addition, for sequences which contain either more or fewer nucleotides than

those of the nucleic acids of the figures, it is understood that the percentage of homology will be determined based on the number of homologous nucleosides in relation to the total number of nucleosides. Thus, for example, homology of sequences shorter than those of the sequences identified herein and as discussed below, will be determined using the number of
5 nucleosides in the shorter sequence.

In one embodiment, the nucleic acid homology is determined through hybridization studies. Thus, *e.g.*, nucleic acids which hybridize under high stringency to a nucleic acid of Tables 1-8, or its complement, or is also found on naturally occurring mRNAs is considered an angiogenesis sequence. In another embodiment, less stringent
10 hybridization conditions are used; for example, moderate or low stringency conditions may be used, as are known in the art; see Ausubel, *supra*, and Tijssen, *supra*.

In addition, the angiogenesis nucleic acid sequences of the invention, *e.g.*, the sequence in Tables 1-8, are fragments of larger genes, *i.e.* they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding
15 and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the angiogenesis genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, *et al.*, *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences, *e.g.*, systems such as
20 UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once the angiogenesis nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire angiogenesis nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, *e.g.*, contained within a plasmid or other vector or excised therefrom as a linear nucleic acid
25 segment, the recombinant angiogenesis nucleic acid can be further-used as a probe to identify and isolate other angiogenesis nucleic acids, for example extended coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant angiogenesis nucleic acids and proteins.

The angiogenesis nucleic acids of the present invention are used in several
30 ways. In a first embodiment, nucleic acid probes to the angiogenesis nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, for example for gene therapy, vaccine, and/or antisense applications. Alternatively, the angiogenesis nucleic acids that include coding regions of angiogenesis

proteins can be put into expression vectors for the expression of angiogenesis proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to angiogenesis nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the angiogenesis nucleic acids, *i.e.* the target sequence (either the target sequence of the sample or to other probe sequences, for example in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally whole genes are not used. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (*i.e.* have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is meant one or more of electrostatic,

hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are
5 attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

10 In general, the probes are attached to the biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified to
15 contain discrete individual sites appropriate for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and
20 copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TeflonJ, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in copending application entitled Reusable
25 Low Fluorescent Plastic Biochip, U.S. Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize
30 sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, for example, the biochip is derivatized with a chemical functional group including, but not

limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, for example using linkers as are known in the art; for example, homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized as is known in the art, and then attached to the surface of the solid support. As will be appreciated by those skilled in the art, either the 5' or 3' terminus may be attached to the solid support, or attachment may be via an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized in situ, using well known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affimetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of angiogenesis-associated sequences. These assays are typically performed in conjunction with reverse transcription. In such assays, an angiogenesis-associated nucleic acid sequence acts as a template in an amplification reaction (*e.g.*, Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of angiogenesis-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, *e.g.*, in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

In some embodiments, a TaqMan based assay is used to measure expression. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent

dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, *e.g.*, AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (*see*, for example, literature provided by Perkin-Elmer, *e.g.*, www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (*see*, Wu and Wallace (1989) *Genomics* 4: 560, Landegren *et al.* (1988) *Science* 241: 1077, and Barringer *et al.* (1990) *Gene* 89: 117), transcription amplification (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173), self-sustained sequence replication (Guatelli *et al.* (1990) *Proc. Nat. Acad. Sci. USA* 87: 1874), dot PCR, and linker adapter PCR, *etc.*

In a preferred embodiment, angiogenesis nucleic acids, *e.g.*, encoding angiogenesis proteins are used to make a variety of expression vectors to express angiogenesis proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (*see, e.g.*, Ausubel, *supra*, and Gene Expression Systems, Fernandez & Hoeffler, Eds, Academic Press, 1999) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the angiogenesis protein. The term "control sequences" refers to DNA sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous,

and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the angiogenesis protein; for example, transcriptional and translational regulatory nucleic acid sequences from *Bacillus* are preferably used to express the angiogenesis protein in *Bacillus*. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

10 In general, transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

15 Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a procaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct.

25 The integrating vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (*e.g.*, Fernandez & Hoeffler, *supra*). See also Kitamura, et al. (1995) PNAS 92:9146-9150.

In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The angiogenesis proteins of the present invention are produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding an angiogenesis protein, under the appropriate conditions to induce or cause expression of the

angiogenesis protein. Conditions appropriate for angiogenesis protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the angiogenesis proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and adenoviral systems. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (see, e.g., Fernandez & Hoeffler, *supra*). Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, angiogenesis proteins are expressed in bacterial systems. Bacterial expression systems are well known in the art. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters

and hybrid promoters are also useful; for example, the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the angiogenesis protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez & Hoeffler, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, angiogenesis proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, angiogenesis protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* and *K. lactis*, *Pichia guilliermondii* and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

The angiogenesis protein may also be made as a fusion protein, using techniques well known in the art. Thus, for example, for the creation of monoclonal antibodies, if the desired epitope is small, the angiogenesis protein may be fused to a carrier protein to form an immunogen. Alternatively, the angiogenesis protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the angiogenesis protein is an angiogenesis peptide, the nucleic acid encoding the peptide may be linked to another nucleic acid for expression purposes. Fusion with detection epitope tags can be made, e.g., with FLAG, His 6, myc, HA, etc.

In one embodiment, the angiogenesis nucleic acids, proteins and antibodies of the invention are labeled. By "labeled" herein is meant that a compound has at least one element, isotope or chemical compound attached to enable the detection of the compound. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) immune labels, which may be antibodies, antigens, or epitope tags and c) colored or fluorescent dyes. The labels may be incorporated into the angiogenesis nucleic acids, proteins and antibodies at any position. For example, the label should be capable of producing, either directly or indirectly, a detectable signal. The detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); David et al., *Biochemistry*, 13:1014 (1974); Pain et al., *J. Immunol. Meth.*, 40:219 (1981); and Nygren, *J. Histochem. and Cytochem.*, 30:407 (1982).

Accordingly, the present invention also provides angiogenesis protein sequences. An angiogenesis protein of the present invention may be identified in several ways. "Protein" in this sense includes proteins, polypeptides, and peptides. As will be appreciated by those in the art, the nucleic acid sequences of the invention can be used to generate protein sequences. There are a variety of ways to do this, including cloning the entire gene and verifying its frame and amino acid sequence, or by comparing it to known sequences to search for homology to provide a frame, assuming the angiogenesis protein has an identifiable motif or homology to some protein in the database being used. Generally, the nucleic acid sequences are input into a program that will search all three frames for homology. This is done in a preferred embodiment using the following NCBI Advanced BLAST parameters. The program is blastx or blastn. The database is nr. The input data is as "Sequence in FASTA format". The organism list is "none". The "expect" is 10; the filter is default. The "descriptions" is 500, the "alignments" is 500, and the "alignment view" is pairwise. The "Query Genetic Codes" is standard (1). The matrix is BLOSUM62; gap existence cost is 11, per residue gap cost is 1; and the lambda ratio is .85 default. This results in the generation of a putative protein sequence.

Also included within one embodiment of angiogenesis proteins are amino acid variants of the naturally occurring sequences, as determined herein. Preferably, the variants are preferably greater than about 75% homologous to the wild-type sequence, more

preferably greater than about 80%, even more preferably greater than about 85% and most preferably greater than 90%. In some embodiments the homology will be as high as about 93 to 95 or 98%. As for nucleic acids, homology in this context means sequence similarity or identity, with identity being preferred. This homology will be determined using standard techniques well known in the art as are outlined above for the nucleic acid homologies.

Angiogenesis proteins of the present invention may be shorter or longer than the wild type amino acid sequences. Thus, in a preferred embodiment, included within the definition of angiogenesis proteins are portions or fragments of the wild type sequences. herein. In addition, as outlined above, the angiogenesis nucleic acids of the invention may be used to obtain additional coding regions, and thus additional protein sequence, using techniques known in the art.

In a preferred embodiment, the angiogenesis proteins are derivative or variant angiogenesis proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative angiogenesis peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at any residue within the angiogenesis peptide.

Also included within one embodiment of angiogenesis proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the angiogenesis protein, using cassette or PCR mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant angiogenesis protein fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the angiogenesis protein amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed angiogenesis variants screened for

the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example, M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants is done using assays of angiogenesis protein activities.

5 Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases deletions may be much larger.

 Substitutions, deletions, insertions or any combination thereof may be used to
10 arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of the angiogenesis protein are desired, substitutions are generally made in accordance with the amino acid substitution chart provided in the definition section.

15 Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those provided in the definition of "conservative substitution". For example, substitutions may be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example the alpha-helical or beta-sheet structure; the charge or hydrophobicity of the
20 molecule at the target site; or the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, *e.g.* seryl or threonyl, is substituted for (or by) a hydrophobic residue, *e.g.* leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain,
25 *e.g.* lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, *e.g.* glutamyl or aspartyl; or (d) a residue having a bulky side chain, *e.g.* phenylalanine, is substituted for (or by) one not having a side chain, *e.g.* glycine.

 The variants typically exhibit the same qualitative biological activity and will elicit the same immune response as the naturally-occurring analog, although variants also are
30 selected to modify the characteristics of the angiogenesis proteins as needed. Alternatively, the variant may be designed such that the biological activity of the angiogenesis protein is altered. For example, glycosylation sites may be altered or removed.

 Covalent modifications of angiogenesis polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino

acid residues of an angiogenesis polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of an angiogenesis polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking angiogenesis polypeptides to a water-insoluble support matrix or surface for use in the method for purifying anti-angiogenesis polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl, threonyl or tyrosyl residues, methylation of the γ -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the angiogenesis polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence angiogenesis polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence angiogenesis polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express angiogenesis-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to angiogenesis polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence angiogenesis polypeptide (for O-linked glycosylation sites). The angiogenesis amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the angiogenesis polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the angiogenesis polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, *CRC Crit. Rev. Biochem.*, pp. 259-306 (1981).

5 Removal of carbohydrate moieties present on the angiogenesis polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., *Arch. Biochem. Biophys.*, 259:52 (1987) and by Edge et al., *Anal. Biochem.*, 118:131
10 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura et al., *Meth. Enzymol.*, 138:350 (1987).

Another type of covalent modification of angiogenesis comprises linking the angiogenesis polypeptide to one of a variety of nonproteinaceous polymers, e.g.,
15 polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

Angiogenesis polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising an angiogenesis polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric
20 molecule comprises a fusion of an angiogenesis polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the angiogenesis polypeptide. The presence of such epitope-tagged forms of an angiogenesis polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the
25 angiogenesis polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of an angiogenesis polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

30 Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 [Field et al., *Mol. Cell. Biol.*, 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., *Molecular and Cellular Biology*, 5:3610-3616 (1985)];

and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [*Paborsky et al., Protein Engineering*, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [*Hopp et al., BioTechnology*, 6:1204-1210 (1988)]; the KT3 epitope peptide [*Martin et al., Science*, 255:192-194 (1992)]; tubulin epitope peptide [*Skinner et al., J. Biol. Chem.*, 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [*Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)].

Also included with an embodiment of angiogenesis protein are other angiogenesis proteins of the angiogenesis family, and angiogenesis proteins from other organisms, which are cloned and expressed as outlined below. Thus, probe or degenerate
10 polymerase chain reaction (PCR) primer sequences may be used to find other related angiogenesis proteins from humans or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include the unique areas of the angiogenesis nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being
15 preferred, and may contain inosine as needed. The conditions for the PCR reaction are well known in the art (*e.g.*, Innis, PCR Protocols, *supra*).

In addition, as is outlined herein, angiogenesis proteins can be made that are longer than those encoded by the nucleic acids of the figures, *e.g.*, by the elucidation of extended sequences, the addition of epitope or purification tags, the addition of other fusion
20 sequences, etc.

Angiogenesis proteins may also be identified as being encoded by angiogenesis nucleic acids. Thus, angiogenesis proteins are encoded by nucleic acids that will hybridize to the sequences of the sequence listings, or their complements, as outlined herein.

25 In a preferred embodiment, when the angiogenesis protein is to be used to generate antibodies, *e.g.*, for immunotherapy or immunodiagnosis, the angiogenesis protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies
30 made to a smaller angiogenesis protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity. In a preferred embodiment, the epitope is selected from a protein sequence set out in Table 8.

Methods of preparing polyclonal antibodies are known to the skilled artisan (e.g., Coligan, *supra*; and Harlow & Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple
5 subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of the figures or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean
10 trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal
15 antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The
20 immunizing agent will typically include a polypeptide encoded by a nucleic acid of Tables 1-8, or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene
25 glycol, to form a hybridoma cell [Goding, *Monoclonal Antibodies: Principles and Practice*, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the
30 growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid Tables 1-8 or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to angiogenesis protein are capable of reducing or eliminating a biological function of an angiogenesis protein, as is described below. That is, the addition of anti-angiogenesis protein antibodies (either polyclonal or preferably monoclonal) to angiogenic tissue (or cells containing angiogenesis) may reduce or eliminate the angiogenesis activity. Generally, at least a 25% decrease in activity is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the angiogenesis proteins are humanized antibodies (*e.g.*, Xenerex Biosciences, Mederex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (*e.g.*, murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues form a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human

immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art.

Generally, a humanized antibody has one or more amino acid residues introduced into it from
5 a source which is non-human. These non-human amino acid residues are often referred to as
import residues, which are typically taken from an import variable domain. Humanization
can be essentially performed following the method of Winter and co-workers [Jones et al.,
Nature, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et
al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the
10 corresponding sequences of a human antibody. Accordingly, such humanized antibodies are
chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact
human variable domain has been substituted by the corresponding sequence from a non-
human species. In practice, humanized antibodies are typically human antibodies in which
some CDR residues and possibly some FR residues are substituted by residues from
15 analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the
art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381
(1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and
Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et
20 al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et
al., *J. Immunol.*, 147(1):86-95 (1991)]. Similarly, human antibodies can be made by
introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the
endogenous immunoglobulin genes have been partially or completely inactivated. Upon
challenge, human antibody production is observed, which closely resembles that seen in
25 humans in all respects, including gene rearrangement, assembly, and antibody repertoire.
This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806;
5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications:
Marks et al., *Bio/Technology* 10, 779-783 (1992); Lonberg et al., *Nature* 368 856-859 (1994);
Morrison, *Nature* 368, 812-13 (1994); Fishwild et al., *Nature Biotechnology* 14, 845-51
30 (1996); Neuberger, *Nature Biotechnology* 14, 826 (1996); Lonberg and Huszar, *Intern. Rev.*
Immunol. 13 65-93 (1995).

By immunotherapy is meant treatment of angiogenesis with an antibody raised
against angiogenesis proteins. As used herein, immunotherapy can be passive or active.
Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient

(patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. As appreciated by one of ordinary skill in the art, the antigen may be provided by injecting a polypeptide against which antibodies are
5 desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the angiogenesis proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory,
10 antibodies used for treatment, bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted angiogenesis protein.

In another preferred embodiment, the angiogenesis protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment, bind the extracellular domain of the angiogenesis protein and prevent it from
15 binding to other proteins, such as circulating ligands or cell-associated molecules. The antibody may cause down-regulation of the transmembrane angiogenesis protein. As will be appreciated by one of ordinary skill in the art, the antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the angiogenesis protein. The antibody is also an antagonist of the angiogenesis protein.
20 Further, the antibody prevents activation of the transmembrane angiogenesis protein. In one aspect, when the antibody prevents the binding of other molecules to the angiogenesis protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to $\text{TNF-}\alpha$, $\text{TNF-}\beta$, IL-1 , $\text{INF-}\gamma$ and IL-2 , or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin,
25 methotrexate, and the like. In some instances the antibody belongs to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, angiogenesis is treated by administering to a patient antibodies directed against the transmembrane angiogenesis protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or
30 otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated or fused to an effector moiety. The effector moiety can be any number of molecules, including labelling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In

one aspect the therapeutic moiety is a small molecule that modulates the activity of the angiogenesis protein. In another aspect the therapeutic moiety modulates the activity of molecules associated with or in close proximity to the angiogenesis protein. The therapeutic moiety may inhibit enzymatic activity such as protease or collagenase activity associated with angiogenesis, or be an attractant of other cells, such as NK cells.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to angiogenesis tissue or cells, results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with angiogenesis. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against angiogenesis proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane angiogenesis proteins not only serves to increase the local concentration of therapeutic moiety in the angiogenesis afflicted area, but also serves to reduce deleterious side effects that may be associated with the therapeutic moiety.

In another preferred embodiment, the angiogenesis protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated or fused to a protein which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the angiogenesis protein can be targeted within a cell, i.e., the nucleus, an antibody thereto contains a signal for that target localization, i.e., a nuclear localization signal.

The angiogenesis antibodies of the invention specifically bind to angiogenesis proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a K_d of at least about 0.1 mM, more usually at least about 1 μ M, preferably at least about 0.1 μ M or better, and most preferably, 0.01 μ M or better. Selectivity of binding is also important.

In a preferred embodiment, the angiogenesis protein is purified or isolated after expression. Angiogenesis proteins may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample. Standard purification methods include electrophoretic, molecular, immunological

and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the angiogenesis protein may be purified using a standard anti-angiogenesis protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For
5 general guidance in suitable purification techniques, see Scopes, R., Protein Purification, Springer-Verlag, NY (1982). The degree of purification necessary will vary depending on the use of the angiogenesis protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the angiogenesis proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection
10 reagents, as vaccine reagents, as screening agents, etc.

Detection of angiogenesis sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the angiogenesis phenotype. Expression levels of genes in normal tissue
15 (*i.e.*, not undergoing angiogenesis) and in angiogenesis tissue (and in some cases, for varying severities of angiogenesis that relate to prognosis, as outlined below) are evaluated to provide expression profiles. An expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state. While two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the
20 generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or angiogenic tissue. This will provide
25 for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, *e.g.*,
30 normal versus angiogenic tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, *e.g.*, in that

expression is increased or decreased; *i.e.*, gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart, Nature Biotechnology, 14:1675-1680 (1996), hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, Northern analysis and RNase protection. As outlined above, preferably the change in expression (*i.e.*, upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript, or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, *e.g.*, with antibodies to the angiogenesis protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to angiogenesis genes, *i.e.*, those identified as being important in an angiogenesis phenotype, can be evaluated in an angiogenesis diagnostic test.

In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In this embodiment, the angiogenesis nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of angiogenesis sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity.

In a preferred embodiment nucleic acids encoding the angiogenesis protein are detected. Although DNA or RNA encoding the angiogenesis protein may be detected, of particular interest are methods wherein an mRNA encoding an angiogenesis protein is detected. Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is

detected. In another method detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding an angiogenesis protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The angiogenesis proteins, antibodies, nucleic acids, modified proteins and cells containing angiogenesis sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, angiogenesis proteins, including intracellular, transmembrane or secreted proteins, find use as markers of angiogenesis. Detection of these proteins in putative angiogenesis tissue allows for detection or diagnosis of angiogenesis. In one embodiment, antibodies are used to detect angiogenesis proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the angiogenesis protein is detected, *e.g.*, by immunoblotting with antibodies raised against the angiogenesis protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the angiogenesis protein find use in *in situ* imaging techniques, *e.g.*, in histology (*e.g.*, *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993)). In this method cells are contacted with from one to many antibodies to the angiogenesis protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the angiogenesis protein(s) contains a detectable label, for example an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and

detectable label. This method finds particular use in simultaneous screening for a plurality of angiogenesis proteins. As will be appreciated by one of ordinary skill in the art, many other histological imaging techniques are also provided by the invention.

5 In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing angiogenesis from biological samples, such as blood, urine, sputum, or other bodily fluids. As previously described, certain angiogenesis proteins are secreted/circulating molecules. 10 Blood samples, therefore, are useful as samples to be probed or tested for the presence of secreted angiogenesis proteins. Antibodies can be used to detect an angiogenesis protein by previously described immunoassay techniques including ELISA, immunoblotting (Western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous angiogenesis protein.

15 In a preferred embodiment, *in situ* hybridization of labeled angiogenesis nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including angiogenesis tissue and/or normal tissue, are made. *In situ* hybridization (*see, e.g., Ausubel, supra*) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the 20 findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the angiogenesis proteins, antibodies, nucleic 25 acids, modified proteins and cells containing angiogenesis sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to angiogenesis severity, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. As above, angiogenesis probes may be attached to biochips for the detection and quantification of angiogenesis sequences in a tissue or 30 patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

In a preferred embodiment members of the three classes of proteins as described herein are used in drug screening assays. The angiogenesis proteins, antibodies, nucleic acids, modified proteins and cells containing angiogenesis sequences are used in drug

screening assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g.,

5 Zlokarnik, et al., *Science* 279, 84-8 (1998); Heid, *Genome Res* 6:986-94, 1996).

In a preferred embodiment, the angiogenesis proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified angiogenesis proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the angiogenesis phenotype or an identified physiological
10 function of an angiogenesis protein. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

15 Having identified the differentially expressed genes herein, a variety of assays may be executed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene as up regulated in angiogenesis, test compounds can be screened for the ability to modulate gene expression or for binding to the angiogenic protein. "Modulation" thus includes both an increase and a decrease in gene
20 expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing angiogenesis, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in angiogenic tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in
25 angiogenic tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the angiogenesis protein and standard
30 immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene expression or protein monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein..

In this embodiment, the angiogenesis nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of angiogenesis sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and
5 analyzed for each well.

Modulators of angiogenesis

Expression monitoring can be performed to identify compounds that modify the expression of one or more angiogenesis-associated sequences, e.g., a polynucleotide sequence set out in Tables 1-8. Generally, in a preferred embodiment, a test modulator is
10 added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate angiogenesis, modulate angiogenesis proteins, bind to an angiogenesis protein, or interfere with the binding of an angiogenesis protein and an antibody or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the angiogenesis phenotype or the expression of an angiogenesis sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter
15 expression profiles, or expression profile nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses an angiogenesis phenotype, for example to a normal tissue fingerprint. In another embodiment, a modulator induced an angiogenesis phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these
20 concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

In one aspect, a modulator will neutralize the effect of an angiogenesis protein. By "neutralize" is meant that activity of a protein is inhibited or blocked and thereby has substantially no effect on a cell.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to an angiogenesis polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property
30

and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop *et al.* (1994) *J. Med. Chem.* 37(9): 1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) *Int. J. Pept. Prot. Res.*, 37: 487-493, Houghton *et al.* (1991) *Nature*, 354: 84-88), peptoids (PCT Publication No WO 91/19735, 26 Dec. 1991), encoded peptides (PCT Publication WO 93/20242, 14 Oct. 1993), random bio-oligomers (PCT Publication WO 92/00091, 9 Jan. 1992), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, (1993) *Proc. Nat. Acad. Sci. USA* 90: 6909-6913), vinyllogous polypeptides (Hagihara *et al.* (1992) *J. Amer. Chem. Soc.* 114: 6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann *et al.*, (1992) *J. Amer. Chem. Soc.* 114: 9217-9218), analogous organic syntheses of small compound libraries (Chen *et al.* (1994) *J. Amer. Chem. Soc.* 116: 2661), oligocarbamates (Cho, et al., (1993) *Science* 261:1303), and/or peptidyl phosphonates (Campbell *et al.*, (1994) *J. Org. Chem.* 59: 658). See, generally, Gordon *et al.*, (1994) *J. Med. Chem.* 37:1385, nucleic acid libraries (see, e.g., Strategene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn *et al.* (1996) *Nature Biotechnology*, 14(3): 309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang *et al.*, (1996) *Science*, 274: 1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum (1993)

C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

5 Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

 A number of well known robotic systems have also been developed for
10 solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. Any of the above devices are suitable for use
15 with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.*, ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences,
20 Columbia, MD, *etc.*).

 The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of angiogenesis gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

25 High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, for example, U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic
30 acid binding (*i.e.*, in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

 In addition, high throughput screening systems are commercially available (*see, e.g.*, Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.*). These systems

typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide
5 detailed protocols for various high throughput systems. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, *e.g.*, cellular extracts containing
10 proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which
15 the target belongs, *e.g.*, substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By
20 "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or
25 most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a
30 limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking;

prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

Modulators of angiogenesis can also be nucleic acids, as defined above.

As described above generally for proteins, nucleic acid modulating agents may
5 be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

10 After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example,
15 an *in vitro* transcription with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, for example, a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme,
20 such as, alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the
25 streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730,
30 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by
5 altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain
10 steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, *e.g.* albumin, detergents, *etc.*
15 which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, *etc.*, may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in
20 expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the angiogenesis phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, *e.g.*, for diagnostic applications, having identified differentially
25 expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the
30 biological activity of the gene product.

In addition screens can be done for genes that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress an angiogenesis expression pattern leading to a normal expression pattern, or to modulate a single angiogenesis gene expression profile so as to mimic the expression of the gene from

normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated angiogenesis tissue reveals genes that are not expressed in normal tissue or angiogenesis tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for angiogenesis genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated angiogenesis tissue sample.

10 Thus, in one embodiment, a test compound is administered to a population of angiogenic cells, that have an associated angiogenesis expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid
15 encoding a proteinaceous candidate agent (*i.e.*, a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, *e.g.*, PCT US97/01019. Regulatable gene therapy systems can also be used.

20 Once the test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

25 Thus, for example, angiogenesis tissue may be screened for agents that modulate, *e.g.*, induce or suppress the angiogenesis phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on angiogenesis activity. By defining such a signature for the angiogenesis phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

30 Measure of angiogenesis polypeptide activity, or of angiogenesis or the angiogenic phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the angiogenesis polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention.

When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of angiogenesis associated with tumors, tumor growth, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (*e.g.*, northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian angiogenesis polypeptide is typically used, *e.g.*, mouse, preferably human.

A variety of angiogenesis assays are known to those of skill in the art. Various models have been employed to evaluate angiogenesis (*e.g.*, Croix *et al.*, *Science* 289:1197-1202, 2000 and Kahn *et al.*, *Amer. J. Pathol.* 156:1887-1900). Assessment of angiogenesis in the presence of a potential modulator of angiogenesis can be performed using cell-culture-based angiogenesis assays, *e.g.*, endothelial cell tube formation assays, as well as other bioassays such as the chick CAM assay, the mouse corneal assay, and assays measuring the effect of administering potential modulators on implanted tumors. The chick CAM assay is described by O'Reilly, *et al.* *Cell* 79: 315-328, 1994. Briefly, 3 day old chicken embryos with intact yolks are separated from the egg and placed in a petri dish. After 3 days of incubation, a methylcellulose disc containing the protein to be tested is applied to the CAM of individual embryos. After about 48 hours of incubation, the embryos and CAMs are observed to determine whether endothelial growth has been inhibited. The mouse corneal assay involves implanting a growth factor-containing pellet, along with another pellet containing the suspected endothelial growth inhibitor, in the cornea of a mouse and observing the pattern of capillaries that are elaborated in the cornea. Angiogenesis can also be measured by determining the extent of neovascularization of a tumor. For example, carcinoma cells can be subcutaneously inoculated into athymic nude mice and tumor growth then monitored. The cancer cells are treated with an angiogenesis inhibitor, such as an antibody, or other compound that is exogenously administered, or can be transfected prior to inoculation with a polynucleotide inhibitor of angiogenesis. Immunoassays using endothelial cell-specific antibodies are typically used to stain for vascularization of tumor and the number of vessels in the tumor.

Assays to identify compounds with modulating activity can be performed *in vitro*. For example, an angiogenesis polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, *e.g.*, from 0.5 to 48 hours. In one embodiment, the angiogenesis polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is measured using immunoassays such as western blotting,

ELISA and the like with an antibody that selectively binds to the angiogenesis polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled
5 detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using the angiogenesis protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or β -gal. The reporter construct is typically transfected into a cell. After
10 treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the
15 expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "angiogenesis proteins". In preferred embodiments the angiogenesis protein comprises a sequence shown in Table 8. The angiogenesis protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

20 Preferably, the angiogenesis protein is a fragment of approximately 14 to 24 amino acids long. More preferably the fragment is a soluble fragment. In one embodiment an angiogenesis protein is conjugated or fused to an immunogenic agent or BSA.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another
25 embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or
30 isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the angiogenesis proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining an angiogenesis protein and a candidate compound, and determining the binding of the compound to the angiogenesis protein. Preferred embodiments utilize the human angiogenesis protein, although other mammalian proteins may also be used, for example for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative angiogenesis proteins may be used.

Generally, in a preferred embodiment of the methods herein, the angiogenesis protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble supports may be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the angiogenesis protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the angiogenesis protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the angiogenesis protein may be done in a number of ways. In a preferred embodiment, the compound is labelled, and binding determined directly, *e.g.*, by attaching all or a portion of the angiogenesis protein to a solid support, adding a labelled candidate agent (*e.g.*, a
5 fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

By "labeled" herein is meant that the compound is either directly or indirectly labeled with a label which provides a detectable signal, *e.g.* radioisotope, fluorescers, enzyme, antibodies, particles such as magnetic particles, chemiluminescers, or specific
10 binding molecules, etc. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin, etc. For the specific binding members, the complementary member would normally be labeled with a molecule which provides for detection, in accordance with known procedures, as outlined above. The label can directly or indirectly provide a detectable signal.

15 In some embodiments, only one of the components is labeled, *e.g.*, the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, *e.g.*, ^{125}I for the proteins and a fluorophore for the compound. Proximity reagents, *e.g.*, quenching or energy transfer reagents are also useful.

20 In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor is a binding moiety known to bind to the target molecule (*i.e.* an angiogenesis protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one
25 embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are typically optimized, *e.g.*, to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally
30 removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by the test compound. Displacement of the competitor is an indication that the test compound is binding to the angiogenesis protein and thus is capable of binding to, and potentially modulating, the

activity of the angiogenesis protein. In this embodiment, either component can be labeled. Thus, for example, if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

5 In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the angiogenesis protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is
10 capable of binding to the angiogenesis protein.

 In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the angiogenesis proteins. In this embodiment, the methods comprise combining an angiogenesis protein and a competitor in a first sample. A second sample comprises a test compound, an angiogenesis protein, and
15 a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the angiogenesis protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the angiogenesis protein.

20 Alternatively, differential screening is used to identify drug candidates that bind to the native angiogenesis protein, but cannot bind to modified angiogenesis proteins. The structure of the angiogenesis protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of an angiogenesis protein are also identified by screening drugs for the ability to either enhance or
25 reduce the activity of the protein.

 Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a
30 radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

 A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, *e.g.* albumin, detergents, *etc.* which may be used

to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

5 In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of an angiogenesis protein. The methods comprise adding a test compound, as defined above, to a cell comprising angiogenesis proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes an angiogenesis protein. In a preferred embodiment, a library of
10 candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, for example hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e. cell-cell contacts). In another
15 example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate angiogenesis agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the angiogenesis protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

20 In one embodiment, a method of inhibiting angiogenic cell division is provided. The method comprises administration of an angiogenesis inhibitor. In another embodiment, a method of inhibiting angiogenesis is provided. The method comprises administration of an angiogenesis inhibitor. In a further embodiment, methods of treating cells or individuals with angiogenesis are provided. The method comprises administration of
25 an angiogenesis inhibitor.

In one embodiment, an angiogenesis inhibitor is an antibody as discussed above. In another embodiment, the angiogenesis inhibitor is an antisense molecule.

Polynucleotide modulators of angiogenesis

30 *Antisense Polynucleotides*

In certain embodiments, the activity of an angiogenesis-associated protein is downregulated, or entirely inhibited, by the use of antisense polynucleotide, *i.e.*, a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, *e.g.*, an angiogenesis protein mRNA, or a subsequence thereof.

Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the angiogenesis protein mRNA. See, *e.g.*, Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, *e.g.*, be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for angiogenesis molecules. A preferred antisense molecule is for an angiogenesis sequences in Tables 1-8, or for a ligand or activator thereof. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of angiogenesis-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (*see, e.g.*, Castanotto *et al.* (1994) *Adv. in Pharmacology* 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, *e.g.*, in Hampel *et al.* (1990) *Nucl. Acids Res.* 18: 299-304; Hampel *et al.* (1990) European Patent Publication No. 0

360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., Wong-Staal *et al.*, WO 94/26877; Ojwang *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90: 6340-6344; Yamada *et al.* (1994) *Human Gene Therapy* 1: 39-45; Leavitt *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92: 699-703; Leavitt *et al.* (1994) *Human Gene Therapy* 5: 1151-120; and Yamada *et al.* (1994) *Virology* 205: 121-126).

Polynucleotide modulators of angiogenesis may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of angiogenesis may be introduced into a cell containing the target nucleic acid sequence, e.g., by formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Thus, in one embodiment, methods of modulating angiogenesis in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-angiogenesis antibody that reduces or eliminates the biological activity of an endogenous angiogenesis protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding an angiogenesis protein. This may be accomplished in any number of ways. In a preferred embodiment, for example when the angiogenesis sequence is down-regulated in angiogenesis, such state may be reversed by increasing the amount of angiogenesis gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous angiogenesis gene or administering a gene encoding the angiogenesis sequence, using known gene-therapy techniques, for example. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), for example as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, for example when the angiogenesis sequence is up-regulated in angiogenesis, the activity of the endogenous angiogenesis gene is decreased, for example by the administration of a angiogenesis antisense nucleic acid or other inhibitor, such as RNAi.

In one embodiment, the angiogenesis eproteins of the present invention may be used to generate polyclonal and monoclonal antibodies to angiogenesis proteins. Similarly, the angiogenesis proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify angiogenesis antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred embodiment, the antibodies are generated to epitopes unique to a angiogenesis protein; that is, the antibodies show little or no cross-reactivity to other proteins. The angiogenesis antibodies may be coupled to standard affinity chromatography columns and used to purify angiogenesis proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the angiogenesis protein.

Methods of identifying variant angiogenesis-associated sequences

Without being bound by theory, expression of various angiogenesis sequences is correlated with angiogenesis. Accordingly, disorders based on mutant or variant angiogenesis genes may be determined. In one embodiment, the invention provides methods for identifying cells containing variant angiogenesis genes, e.g., determining all or part of the sequence of at least one endogeneous angiogenesis genes in a cell. This may be accomplished using any number of sequencing techniques. In a preferred embodiment, the invention provides methods of identifying the angiogenesis genotype of an individual, e.g., determining all or part of the sequence of at least one angiogenesis gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced angiogenesis gene to a known angiogenesis gene, i.e., a wild-type gene.

The sequence of all or part of the angiogenesis gene can then be compared to the sequence of a known angiogenesis gene to determine if any differences exist. This can be done using any number of known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the angiogenesis gene of the patient and the known angiogenesis gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the angiogenesis genes are used as probes to determine the number of copies of the angiogenesis gene in the genome.

In another preferred embodiment, the angiogenesis genes are used as probes to determine the chromosomal localization of the angiogenesis genes. Information such as

chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the angiogenesis gene locus.

5 *Administration of pharmaceutical and vaccine compositions*

In one embodiment, a therapeutically effective dose of an angiogenesis protein or modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using
10 known techniques (*e.g.*, Ansel *et al.*, *Pharmaceutical Dosage Forms and Drug Delivery*, Lippincott, Williams & Wilkins Publishers, ISBN:0683305727; Lieberman (1992) *Pharmaceutical Dosage Forms* (vols. 1-3), Dekker, ISBN 0824770846, 082476918X, 0824712692, 0824716981; Lloyd (1999) *The Art, Science and Technology of Pharmaceutical Compounding*, Amer. Pharmaceutical Assn, ISBN 0917330889; and Pickar (1999) *Dosage*
15 *Calculations*, Delmar Pub, ISBN 0766805042). As is known in the art, adjustments for angiogenesis degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration, drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art.

20 A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the angiogenesis proteins and modulators thereof of the
25 present invention can be done in a variety of ways as discussed above, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, for example, in the treatment of wounds and inflammation, the angiogenesis proteins and modulators may be directly applied as a solution or spray.

30 The pharmaceutical compositions of the present invention comprise an angiogenesis protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain

the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that angiogenesis protein modulators (*e.g.*, antibodies, antisense constructs, ribozymes, small organic molecules, *etc.*) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise an angiogenesis protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, *e.g.*, buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents

and the like, for example, sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania (1980) and Goodman and Gillman, *The Pharmacological Basis of Therapeutics*, (Hardman, J.G, Limbird, L.E, Molinoff, P.B., Ruddon, R.W, and Gilman, A.G., eds) The McGraw-Hill Companies, Inc., 1996).

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art, e.g., *Remington's Pharmaceutical Science* and Goodman and Gillman, *The Pharmacological Basis of Therapeutics*, *supra*.

The compositions containing modulators of angiogenesis proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer.

It will be appreciated that the present angiogenesis protein-modulating compounds can be administered alone or in combination with additional angiogenesis modulating compounds or with other therapeutic agent, *e.g.*, other anti-cancer agents or treatments.

5 In numerous embodiments, one or more nucleic acids, *e.g.*, polynucleotides comprising nucleic acid sequences set forth in Tables 1-8, such as antisense polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of angiogenesis-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo* or *in vivo* (cell or
10 organism-based) recombinant expression systems.

 The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection,
15 plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* volume 152 Academic Press, Inc., San Diego, CA (Berger), F.M. Ausubel *et al.*, eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley &
20 Sons, Inc., (supplemented through 1999), and Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.

 In a preferred embodiment, angiogenesis proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly,
25 angiogenesis genes (including both the full-length sequence, partial sequences, or regulatory sequences of the angiogenesis coding regions) can be administered in a gene therapy application. These angiogenesis genes can include antisense applications, either as gene therapy (*i.e.* for incorporation into the genome) or as antisense compositions, as will be appreciated by those in the art.

30 Angiogenesis polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL and antibody responses.. Such vaccine compositions can include, for example, lipidated peptides (*e.g.*, Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991; Alonso

- et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995), peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998), multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996), peptides formulated as
- 5 multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F.
- 10 H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990), particles of viral or synthetic origin (*e.g.*, Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993), liposomes (Reddy, R. *et al.*, *J.*
- 15 *Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996), or, naked or particle absorbed cDNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993).
- 20 Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or

25 *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of

30 acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff *et al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; 5 WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the 10 invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, for example, as a vector to express nucleotide sequences that encode angiogenic polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an 15 immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization e.g. adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified 20 anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein (see, e.g., Shata *et al.* (2000) *Mol Med Today*, 6: 66-71; Shedlock *et al.*, *J Leukoc Biol* 68,:793-806, 2000; Hipp *et al.*, *In Vivo* 14:571-85, 2000).

Methods for the use of genes as DNA vaccines are well known, and include placing an angiogenesis gene or portion of an angiogenesis gene under the control of a 25 regulatable promoter or a tissue-specific promoter for expression in an angiogenesis patient. The angiogenesis gene used for DNA vaccines can encode full-length angiogenesis proteins, but more preferably encodes portions of the angiogenesis proteins including peptides derived from the angiogenesis protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from an angiogenesis gene. 30 For example, angiogenesis-associated genes or sequence encoding subfragments of an angiogenesis protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, the DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the angiogenesis polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

5 In another preferred embodiment angiogenesis genes find use in generating animal models of angiogenesis. When the angiogenesis gene identified is repressed or diminished in angiogenic tissue, gene therapy technology, *e.g.*, wherein antisense RNA directed to the angiogenesis gene will also diminish or repress expression of the gene. Animal models of angiogenesis find use in screening for modulators of an angiogenesis-associated sequence or modulators of angiogenesis. Similarly, transgenic animal technology
10 including gene knockout technology, for example as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the angiogenesis protein. When desired, tissue-specific expression or knockout of the angiogenesis protein may be necessary.

15 It is also possible that the angiogenesis protein is overexpressed in angiogenesis. As such, transgenic animals can be generated that overexpress the angiogenesis protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression
20 level of the transgene. Animals generated by such methods find use as animal models of angiogenesis and are additionally useful in screening for modulators to treat angiogenesis or to evaluate a therapeutic entity.

Kits for Use in Diagnostic and/or Prognostic Applications

25 For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In the diagnostic and research applications such kits may include any or all of the following: assay reagents, buffers, angiogenesis-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, dominant negative angiogenesis polypeptides or polynucleotides, small molecules
30 inhibitors of angiogenesis-associated sequences *etc.* A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing directions (*i.e.*, protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. Any

medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

The present invention also provides for kits for screening for modulators of angiogenesis-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: an angiogenesis-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing angiogenic-associated activity. Optionally, the kit contains biologically active angiogenesis protein. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

EXAMPLES

Example 1: Tissue Preparation, Labeling Chips, and Fingerprints

Purify total RNA from tissue using TRIzol Reagent

Homogenize tissue samples in 1ml of TRIzol per 50mg of tissue using a Polytron 3100 homogenizer. The generator/probe used depends upon the tissue size. A generator that is too large for the amount of tissue to be homogenized will cause a loss of sample and lower RNA yield. TRIzol is added directly to frozen tissue, which is then homogenize. Following homogenization, insoluble material is removed by centrifugation at 7500 x g for 15 min in a Sorvall superspeed or 12,000 x g for 10 min. in an Eppendorf centrifuge at 4°C. The clear homogenate is transferred to a new tube for use. The samples may be frozen now at -60° to -70°C (and kept for at least one month). The homogenate is

mixed with 0.2ml of chloroform per 1ml of TRIzol reagent used in the original homogenization and incubated at room temp. for 2-3 minutes. The aqueous phase is then separated by centrifugation and transferred to a fresh tube and the RNA precipitated using isopropyl alcohol. The pellet is isolated by centrifugation, washed, air-dried, resuspended in
5 an appropriate volume of DEPC H₂O, and the absorbance measured.

Purification of poly A+ mRNA from total RNA is performed as follows. Heat an oligotex suspension to 37°C and mixing immediately before adding to RNA. The Elution Buffer is heated at 70°C. Warm up 2 x Binding Buffer at 65°C if there is precipitate in the buffer. Mix total RNA with DEPC-treated water, 2 x Binding Buffer, and Oligotex
10 according to Table 2 on page 16 of the Oligotex Handbook. Incubate for 3 minutes at 65°C. Incubate for 10 minutes at room temperature. Centrifuge for 2 minutes at 14,000 to 18,000 g. Remove supernatant without disturbing Oligotex pellet. A little bit of solution can be left behind to reduce the loss of Oligotex. Gently resuspend in Wash Buffer OW2 and pipet onto spin column. Centrifuge the spin column at full speed for 1 minute. Transfer spin column to
15 a new collection tube and gently resuspend in Wash Buffer OW2 and centrifuge as describe herein. Transfer spin column to a new tube and elute with 20 to 100 ul of preheated (70°C) Elution Buffer. Gently resuspend Oligotex resin by pipetting up and down. Centrifuge as above. Repeat elution with fresh elution buffer or use first eluate to keep the elution volume low. Read absorbance, using diluted Elution Buffer as the blank. Before proceeding with
20 cDNA synthesis, precipitate the mRNA as follows: add 0.4 vol. of 7.5 M NH₄OAc + 2.5 vol. of cold 100% ethanol. Precipitate at -20°C 1 hour to overnight (or 20-30 min. at -70°C). Centrifuge at 14,000-16,000 x g for 30 minutes at 4°C. Wash pellet with 0.5ml of 80% ethanol (-20°C) then centrifuge at 14,000-16,000 x g for 5 minutes at room temperature.. Repeat 80% ethanol wash. Air dry the ethanol from the pellet in the hood.. Suspend pellet in
25 DEPC H₂O at 1ug/ul concentration.

To further Clean up total RNA using Qiagen's RNeasy kit, add no more than 100ug to an RNeasy column. Adjust sample to a volume of 100ul with RNase-free water. Add 350ul Buffer RLT then 250ul ethanol (100%) to the sample. Mix by pipetting (do not centrifuge) then apply sample to an RNeasy mini spin column. Centrifuge for 15 sec at
30 >10,000rpm. Transfer column to a new 2-ml collection tube. Add 500ul Buffer RPE and centrifuge for 15 sec at >10,000rpm. Discard flowthrough. Add 500ul Buffer RPE and centrifuge for 15 sec at >10,000rpm. Discard flowthrough then centrifuge for 2 min at maximum speed to dry column membrane. Transfer column to a new 1.5-ml collection tube

and apply 30-50ul of RNase-free water directly onto column membrane. Centrifuge 1 min at >10,000rpm. Repeat elution. and read absorbance.

cDNA synthesis using Gibco's "SuperScript Choice System for cDNA Synthesis" kit

- 5 First Strand cDNA synthesis is performed as follows. Use 5ug of total RNA or 1ug of polyA+ mRNA as starting material. For total RNA, use 2ul of SuperScript RT. For polyA+ mRNA, use 1ul of SuperScript RT. Final volume of first strand synthesis mix is 20ul. RNA must be in a volume no greater than 10ul. Incubate RNA with 1ul of 100pmol T7-T24 oligo for 10 min at 70C. On ice, add 7 ul of: 4ul 5X 1st Strand Buffer, 2ul of 0.1M DTT, and 1 ul of 10mM dNTP mix. Incubate at 37C for 2 min then add SuperScript RT. Incubate at 37C for 1 hour.

- For the second strand synthesis, place 1st strand reactions on ice and add: 91ul DEPC H₂O; 30ul 5X 2nd Strand Buffer; 3ul 10mM dNTP mix; 1ul 10U/ul E.coli DNA Ligase; 4ul 10U/ul E.coli DNA Polymerase; and 1ul 2U/ul RNase H. Mix and incubate 2 hours at 16C. Add 2ul T4 DNA Polymerase. Incubate 5 min at 16C. Add 10ul of 0.5M EDTA. A further clean-up of DNA is performed using phenol:chloroform:isoamyl Alcohol (25:24:1) purification.

- In vitro* Transcription (IVT) and labeling with biotin is performed as follows: Pipet 1.5ul of cDNA into a thin-wall PCR tube. Make NTP labeling mix by combining 2ul T7 10xATP (75mM) (Ambion); 2ul T7 10xGTP (75mM) (Ambion); 1.5ul T7 10xCTP (75mM) (Ambion); 1.5ul T7 10xUTP (75mM) (Ambion); 3.75ul 10mM Bio-11-UTP (Boehringer-Mannheim/Roche or Enzo); 3.75ul 10mM Bio-16-CTP (Enzo); 2ul 10x T7 transcription buffer (Ambion); and 2ul 10x T7 enzyme mix (Ambion). The final volume is 20ul. Incubate 6 hours at 37°C in a PCR machine. The RNA can be further cleaned.

- 25 Fragmentation is performed as follows. 15 ug of labeled RNA is usually fragmented. Try to minimize the fragmentation reaction volume; a 10 ul volume is recommended but 20 ul is all right. Do not go higher than 20 ul because the magnesium in the fragmentation buffer contributes to precipitation in the hybridization buffer. Fragment RNA by incubation at 94 C for 35 minutes in 1 x Fragmentation buffer (5 x Fragmentation buffer is 200 mM Tris-acetate, pH 8.1; 500 mM KOAc; 150 mM MgOAc). The labeled RNA transcript can be analyzed before and after fragmentation. Samples can be heated to 65°C for 15 minutes and electrophoresed on 1% agarose/TBE gels to get an approximate idea of the transcript size range

For hybridization, 200 μ l (10 μ g cRNA) of a hybridization mix is put on the chip. If multiple hybridizations are to be done (such as cycling through a 5 chip set), then it is recommended that an initial hybridization mix of 300 μ l or more be made. The hybridization mix is: fragment labeled RNA (50ng/ μ l final conc.); 50 pM 948-b control
 5 oligo; 1.5 pM BioB; 5 pM BioC; 25 pM BioD; 100 pM CRE; 0.1mg/ml herring sperm DNA; 0.5mg/ml acetylated BSA; and 300 μ l with 1xMES hyb buffer.

Labeling is performed as follows: The hybridization reaction includes non-biotinylated IVT (purified by RNeasy columns); IVT antisense RNA 4 μ g/ μ l; random Hexamers (1 μ g/ μ l) 4 μ l and water to 14 μ l. The reaction is incubated at 70°C, 10 min.
 10 Reverse transcription is performed in the following reaction: 5X First Strand (BRL) buffer, 6 μ l; 0.1 M DTT, 3 μ l; 50X dNTP mix, 0.6 μ l; H₂O, 2.4 μ l; Cy3 or Cy5 dUTP (1mM), 3 μ l; SS RT II (BRL), 1 μ l in a final volume of 16 μ l. Add to hybridization reaction. Incubate 30 min., 42°C. Add 1 μ l SSII and incubate another hour. Put on ice. 50X dNTP mix (25mM of cold dATP, dCTP, and dGTP, 10mM of dTTP: 25 μ l each of 100mM dATP, dCTP, and
 15 dGTP; 10 μ l of 100mM dTTP to 15 μ l H₂O. dNTPs from Pharmacia)

RNA degradation is performed as follows. Add 86 μ l H₂O, 1.5 μ l 1M NaOH/2mM EDTA and incubate at 65°C, 10 min.. For U-Con 30, 500 μ l TE/sample spin at 7000g for 10 min, save flow through for purification. For Qiagen purification, suspend u-con recovered material in 500 μ l buffer PB and proceed using Qiagen protocol. For DNase
 20 digestion, add 1 μ l of 1/100 dil of DNase/30 μ l Rx and incubate at 37°C for 15 min. Incubate at 5 min 95°C to denature the DNase/

For sample preparation, add Cot-1 DNA, 10 μ l; 50X dNTPs, 1 μ l; 20X SSC, 2.3 μ l; Na pyro phosphate, 7.5 μ l; 10mg/ml Herring sperm DNA; 1 μ l of 1/10 dilution to 21.8 final vol. Dry in speed vac. Resuspend in 15 μ l H₂O. Add 0.38 μ l 10% SDS. Heat 95°C, 2
 25 min and slow cool at room temp. for 20 min. Put on slide and hybridize overnight at 64°C. Washing after the hybridization: 3X SSC/0.03% SDS: 2 min., 37.5 mls 20X SSC+0.75mls 10% SDS in 250mls H₂O; 1X SSC: 5 min., 12.5 mls 20X SSC in 250mls H₂O; 0.2X SSC: 5 min., 2.5 mls 20X SSC in 250mls H₂O. Dry slides and scan at appropriate PMT's and channels.

30

Example 2. A model of angiogenesis is used to determine expression in angiogenesis

In the model of angiogenesis used to determine expression of angiogenesis-associated sequences, human umbilical vein endothelial cells (HUVEC) were obtained, *e.g.*,

as passage 1 (p1) frozen cells from Cascade Biologics (Oregon) and grown in maintenance medium: Medium 199 (Life Technologies) supplemented with 20% pooled human serum, 100 mg/ml heparin and 75 mg/ml endothelial cell growth supplements (Sigma) and gentamicin (Life Technologies). An *in vitro* cell system model was used in which 2×10^5 5 HUVECs were cultured in 0.5 ml 3 mgs/ml plasminogen-depleted fibrinogen (Calbiochem, San Diego, CA) that was polymerized by the addition of 1 unit of maintenance medium supplemented with 100 ng/ml VEGF and HGF and 10 ng/ml TGF- α (R&D Systems, Minneapolis, MN) added (growth medium). The growth medium was replaced every 2 days. Samples for RNA were collected, *e.g.*, at 0, 2, 6, 15, 24, 48, and 96 hours of culture. The 10 fibrin clots were placed in Trizol (Life Technologies) and disrupted using a Tissueomizer. Thereafter standard procedures were used for extracting the RNA (*e.g.*, Example 1).

Angiogenesis associated sequences thus identified are shown in Tables 1-8 . As indicated, some of the Accession numbers include expression sequence tags (ESTs). Thus, in one embodiment herein, genes within an expression profile, also termed expression 15 profile genes, include ESTs and are not necessarily full length.

TABLE 1:

5	Pkey:	Unique Eos probeset identifier number			
	Accession:	Accession number used for previous patent filings			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
10	Pkey	Accession	ExAccn	UnigeneID	UnigeneTitle
15	134404	AB000450	AB000450	Hs.82771	vaccinia related kinase 2
	121443	AB002380	AF180681	Hs.6582	Rho guanine exchange factor (GEF) 12
	100082	AB003103	AA130080	Hs.4295	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12
	132817	AB004884	N27852	Hs.57553	tousled-like kinase 2
	130150	AF000573_ma1	BE094848	Hs.15113	homogentisate 1,2-dioxygenase (homogentisate oxidase)
	100104	AF008937	AF008937	Hs.102178	syntaxin 16
20	130839	AF009301	AB011169	Hs.20141	similar to S. cerevisiae SSM4
	427064	AF009368	AF029674	Hs.173422	KIAA1605 protein
	100113	D00591	NM_001269	Hs.84746	chromosome condensation 1
	133980	D00760	AA294921	Hs.250811	v-rat simian leukemia viral oncogene homolog B (ras related; GTP binding protein)
	100129	D11139	AA469369	Hs.5831	tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)
25	100154	D14657	H60720	Hs.81892	KIAA0101 gene product
	100169	D14878	AL037228	Hs.82043	D123 gene product
	101956	D17716	NM_002410	Hs.121502	mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase
	100190	D21090	M91401	Hs.178658	RAD23 (S. cerevisiae) homolog B
	134742	D26135	NM_001346	Hs.89462	diacylglycerol kinase, gamma (90kD)
30	100211	D26528	D26528	Hs.123058	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 7 (RNA helicase, 52kD)
	100238	D30742	L24959	Hs.348	calcium/calmodulin-dependent protein kinase IV
	130283	D31762	NM_012288	Hs.153954	TRAM-like protein
	134237	D31765	D31765	Hs.170114	KIAA0061 protein
	100248	D31888	NM_015156	Hs.78398	KIAA0071 protein
35	100256	D38128	D25418	Hs.393	prostaglandin I2 (prostaglandin) receptor (IP)
	100262	D38500	D38500	Hs.278468	postmeiotic segregation Increased 2-like 4
	134329	D38551	N92036	Hs.81848	RAD21 (S. pombe) homolog
	100281	D42087	AF091035	Hs.184627	KIAA0118 protein
	100294	D49396	AA331881	Hs.75454	peroxiredoxin 3
40	100327	D55640	D55640		gb:human monocyte PABL (pseudautosomal boundary-like sequence) mRNA, clone Mo2.
	100335	D63391	AW247529	Hs.6793	platelet-activating factor acetylhydrolase, isoform 1b, gamma subunit (29kD)
	134495	D63477	D63477	Hs.84087	KIAA0143 protein
	100338	D63483	D66864	Hs.57735	acetyl LDL receptor, SREC
	135152	D64015	M96954	Hs.182741	TIA1 cytotoxic granule-associated RNA-binding protein-like 1
45	134269	D79990	NM_014737	Hs.80905	Ras association (RalGDS/AF-6) domain family 2
	100372	D79997	NM_014791	Hs.184339	KIAA0175 gene product
	134304	D80010	BE613486	Hs.81412	lipin 1
	100394	D84276	D84284	Hs.66052	CD38 antigen (p45)
	100405	D86425	AW291587	Hs.82733	nidogen 2
50	100418	D86978	D86978	Hs.84790	KIAA0225 protein
	133154	D87012	D87012	Hs.194685	topoisomerase (DNA) III beta
	134347	D87075	AF164142	Hs.82042	solute carrier family 23 (nucleobase transporters), member 1
	128653	D87432	D87432	Hs.10315	solute carrier family 7 (cationic amino acid transporter, y+ system), member 6
	100438	D87448	AA013051	Hs.91417	topoisomerase (DNA) II binding protein
55	134593	D87845	NM_000437	Hs.234392	platelet-activating factor acetylhydrolase 2 (40kD)
	100481	HG1098-HT1098	X70377	Hs.121489	cystatin D
	100552	HG2167-HT2237	AA019521	Hs.301946	lysosomal
	100591	HG2415-HT2511	NM_004091	Hs.231444	Homo sapiens, Similar to hypothetical protein PRO1722, clone MGC:15692, mRNA, complete cds
60	100652	HG2825-HT2949	BE613608	Hs.142653	rat finger protein
	100662	HG2887-HT3031_r	AI368680	Hs.816	SRY (sex determining region Y)-box 2
	100899	HG4660-HT5073	AL039123	Hs.103042	microtubule-associated protein 1B
	100905	HG4704-HT5146	L12260	Hs.172816	neuregulin 1
	100945	HG884-HT884	AF002225	Hs.180686	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)
65	100950	HG919-HT919	AF128542	Hs.166846	polymerase (DNA directed), epsilon
	100964	J00212_f	J00212		Empirically selected from AFFX single probeset
	135407	J04029	J04029	Hs.99936	keratin 10 (epidermolytic hyperkeratosis; keratosis palmaris et plantaris)
	130149	J04031	AW067805	Hs.172665	methylene-tetrahydrofolate dehydrogenase (NADP+ dependent), methylenetetrahydrofolate
	131877	J04088	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)
	101016	J04543	J04543	Hs.78637	annexin A7
70	134786	L06139	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous malformations, multiple cutaneous and mucosal)
	134100	L07540	AA460085	Hs.171075	replication factor C (activator 1) 5 (36.5kD)
	134078	L08895	L08895	Hs.78895	MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C)
	101132	L11239	L11239	Hs.36993	gastrulation brain homeo box 1
	134849	L11353	BE409525	Hs.902	neurofibromin 2 (bilateral acoustic neuroma)
75	106432	L13773	AK000310	Hs.17138	hypothetical protein FLJ20303

5	101152	L13800	A1884625	Hs.9884	spindle pole body protein
	135397	L14922	L14822	Hs.166563	replication factor C (activator 1) 1 (145kD)
	131687	L15189	BE297635	Hs.3089	heat shock 70kD protein 9B (mortalin-2)
	101168	L15388	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	421155	L16895	H87879	Hs.102267	lysyl oxidase
10	101226	L27476	AF083892	Hs.75608	tight junction protein 2 (zona occludens 2)
	133975	L27624	C18356	Hs.295944	tissue factor pathway inhibitor 2
	134739	L32976	NM_002419	Hs.89449	mitogen-activated protein kinase kinase kinase 11
	130155	L33404	AA101043	Hs.151254	kallikrein 7 (chymotryptic, stratum corneum)
	440538	L35263	W76332	Hs.79107	mitogen-activated protein kinase 14
15	132813	L37347	BE313625	Hs.57435	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2
	101294	L40371	AF168418	Hs.116784	thyroid hormone receptor interactor 4
	101300	L40391	BE535511	Hs.74137	transmembrane trafficking protein
	101310	L41607	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2, l-branching enzyme
	130344	L77566	AW250122	Hs.154879	DiGeorge syndrome critical region gene DGS1; likely ortholog of mouse expressed sequence 2
20		embryonic lethal			
	101381	M13928	AW675039	Hs.1227	antholevullinate, delta-, dehydratase
	101688	M14016	AW005903	Hs.78501	uroporphyrinogen decarboxylase
	133780	M14219	AA557660	Hs.76152	decorin
	101396	M15796	BE267931	Hs.78996	proliferating cell nuclear antigen
25	101447	M21305	M21305		gbHuman alpha satellite and satellite 3 junction DNA sequence.
	101458	M22092	M22092		gbHuman neural cell adhesion molecule (N-CAM) gene, exon SEC and partial cds.
	101470	M22898	NM_000546	Hs.1846	tumor protein p53 (LI-Fraumeni syndrome)
	134604	M22995	NM_002884	Hs.865	RAP1A, member of RAS oncogene family
	101478	M23379	NM_002890	Hs.758	RAS p21 protein activator (GTPase activating protein) 1
30	406698	M24364	X03068	Hs.73931	major histocompatibility complex, class II, DQ beta 1
	133519	M24400	AW583062	Hs.74502	chymotrypsinogen B1
	131185	M25753	BE280074	Hs.23960	cyclin B1
	134116	M27691	R84694	Hs.79194	cAMP responsive element binding protein 1
	133999	M28213	AA535244	Hs.78305	RAB2, member RAS oncogene family
35	130174	M29550	M29551	Hs.151531	protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform (calcineurin A beta)
	129963	M29971	M29971	Hs.1384	O-6-methylguanine-DNA methyltransferase
	132983	M30269	M30269	Hs.62041	nkdog (enactin)
	133900	M31158	M31158	Hs.77439	protein kinase, cAMP-dependent, regulatory, type II, beta
	101543	M31166	M31166	Hs.2050	pentactin-related gene, rapidly induced by IL-1 beta
40	101545	M31210	BE246154	Hs.154210	endothelial differentiation, sphingolipid G-protein-coupled receptor, 1
	101620	M56420	S55271	Hs.247930	Epsilon . IgE
	134691	M59979	AW382987	Hs.88474	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
	133595	M62810	AA393273	Hs.75133	transcription factor 6-like 1 (mitochondrial transcription factor 1-like)
	130425	M63383	AA243383	Hs.155530	interferon, gamma-inducible protein 16
45	101700	M64710	D90337	Hs.247916	natriuretic peptide precursor C
	101714	M68874	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic, calcium-dependent)
	134246	M74524	D28459	Hs.80612	ubiquitin-conjugating enzyme E2A (RAD6 homolog)
	101760	M80254	M80254	Hs.173125	peptidylprolyl isomerase F (cyclophilin F)
	133948	M81780_cds3	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid lysosomal (acid sphingomyelinase)
50	101791	M83822	M83822	Hs.62354	cell division cycle 4-like
	101812	M86934	BE439894	Hs.78991	DNA segment, numerous copies, expressed probes (GS1 gene)
	101813	M87338	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40kD)
	133396	M96326_ma1	M96326	Hs.72885	azurocidin 1 (cationic antimicrobial protein 37)
	135152	M96954	M96954	Hs.182741	TIA1 cytotoxic granule-associated RNA-binding protein-like 1
55	129026	M98833	AL120297	Hs.108043	Friend leukemia virus integration 1
	101901	S66793	H38026	Hs.308	arrestin 3, retinal (X-arrestin)
	134831	S72370	AA853479	Hs.89890	pyruvate carboxylase
	134039	S78569	NM_002290	Hs.78672	laminin, alpha 4
	134395	S79873	AA456539	Hs.8262	lysosomal
60	101975	S83325	AA079717	Hs.283664	aspartate beta-hydroxylase
	101977	S83364	AF112213	Hs.184062	putative Rab5-interacting protein
	101978	S83365	BE561610	Hs.5809	putative transmembrane protein; homolog of yeast Golgi membrane protein Yif1p (Yip1p-interacting factor)
	101998	U01212	U01212	Hs.248153	olfactory marker protein
	102003	U01922	U01922	Hs.125565	translocase of inner mitochondrial membrane 8 (yeast) homolog A
65	102007	U02556	U02556	Hs.75307	t-complex-associated-testis-expressed 1-like
	102009	U02680	BE245149	Hs.82643	protein tyrosine kinase 9
	416658	U03272	U03272	Hs.79432	fibrillin 2 (congenital contractual arachnodactyly)
	132951	U04209	AW821182	Hs.51418	microfibrillar-associated protein 1
	135389	U05237	U05237	Hs.99872	fetal Alzheimer antigen
70	102048	U07225	U07225	Hs.339	purinergic receptor P2Y, G-protein coupled, 2
	130145	U07620	U34820	Hs.151051	mitogen-activated protein kinase 10
	303153	U09759	U09759	Hs.246857	mitogen-activated protein kinase 9
	420269	U09820	U72937	Hs.98264	alpha thalassemia/mental retardation syndrome X-linked (RAD54 (S. cerevisiae) homolog)
	102095	U11313	U11313	Hs.75760	sterol carrier protein 2
75	102123	U14518	NM_001809	Hs.1594	centromere protein A (17kD)
	102126	U14575	AW950870	Hs.78961	protein phosphatase 1, regulatory (inhibitor) subunit 8
	102133	U15173	AU076845	Hs.155596	BCL2/adenovirus E1B 19kD-interacting protein 2
	102139	U15932	NM_004419	Hs.2128	dual specificity phosphatase 5
	102162	U18291	AA450274	Hs.1592	CDC16 (cell division cycle 16, S. cerevisiae, homolog)

	102164	U18300	NM_000107	Hs.77602	damage-specific DNA binding protein 2 (48kD)
	427653	U18383	AA159001	Hs.180069	nuclear respiratory factor 1
	131817	U20536	U20536	Hs.3280	caspase 6, apoptosis-related cysteine protease
	102200	U21551	AA232362	Hs.157205	branched chain aminotransferase 1, cytosolic
5	102210	U23028	BE619413	Hs.2437	eukaryotic translation initiation factor 2B, subunit 5 (epsilon, 82kD)
	102214	U23752	U23752	Hs.32964	SRY (sex determining region Y)-box 11
	132811	U25435	U25435	Hs.57419	CCCTC-binding factor (zinc finger protein)
	131319	U25997	NM_003155	Hs.25590	stanniocalcin 1
10	102256	U28251_cds2	U28251	Hs.53237	ESTs, Highly similar to Z169_HUMAN ZINC FINGER PROTEIN 169 [H.sapiens]
	132316	U28831	U28831	Hs.44566	KIAA1641 protein
	102269	U30245	U30245		gb:human myelomonocytic specific protein (MND) gene, 5' flanking sequence and complete exon 1.
	134365	U32315	AA568906	Hs.82240	synaptobin 3A
	102293	U32439	AF090116	Hs.79348	regulator of G-protein signaling 7
15	102298	U32849	AA382169	Hs.54483	N-myc (and STAT) Interactor
	102325	U35139	AI815887	Hs.50130	neodin (mouse) homolog
	302344	U36764	BE303044	Hs.192023	eukaryotic translation initiation factor 3, subunit 2 (beta, 36kD)
	102361	U39400	AA223616	Hs.75859	chromosome 11 open reading frame 4
	102367	U39657	U39656	Hs.118825	mitogen-activated protein kinase kinase 6
20	102388	U41344	AA362907	Hs.78494	proline arginine-rich end leucine-rich repeat protein
	102394	U41766	NM_003816	Hs.2442	a disintegrin and metalloproteinase domain 9 (metrin gamma)
	128829	U41813	AF010258	Hs.127428	homeo box A9
	102251	U41815	NM_004398	Hs.41706	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 10 (RNA helicase)
	102409	U43286	BE300330	Hs.118725	selenophosphate synthetase 2
25	133746	U44378	AW410035	Hs.75862	MAD (mothers against decapentaplegic, Drosophila) homolog 4
	102423	U44754	Z47542	Hs.179312	small nuclear RNA activating complex, polypeptide 1, 43kD
	132828	U47011_cds1	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-induced)
	130441	U47077	U63630	Hs.155637	protein kinase, DNA-activated, catalytic polypeptide
	102450	U48251	U48251	Hs.75871	protein kinase C binding protein 1
30	129350	U50535	U50535	Hs.110630	Human BRCA2 region, mRNA sequence CG006
	102534	U56833	U96759	Hs.198307	von Hippel-Lindau binding protein 1
	130457	U58091	AB014595	Hs.155976	cutlin 4B
	135065	U58837	AA019401	Hs.93909	cyclic nucleotide gated channel beta 1
	102560	U59289	R97457	Hs.63984	cadherin 13, H-cadherin (heart)
35	102567	U59863	U63830	Hs.146847	TRAF family member-associated NFKB activator
	134305	U67122	U61397	Hs.81424	ubiquitin-like 1 (sentrin)
	102638	U67319	U67319	Hs.9216	caspase 7, apoptosis-related cysteine protease
	132736	U68019	AW081883	Hs.288261	Homo sapiens cDNA: FLJ23037 fls, clone LNG02036, highly similar to HSU68019 Homo
40					sapiens mad protein homolog (hMAD-3) mRNA
	133070	U69611	U92649	Hs.64311	a disintegrin and metalloproteinase domain 17 (tumor necrosis factor, alpha, converting enzyme)
	102663	U70322	NM_002270	Hs.168075	karyopherin (importin) beta 2
	134660	U73524	U73524	Hs.87465	ATP/GTP-binding protein
	102735	U79267	AF111106	Hs.3382	protein phosphatase 4, regulatory subunit 1
	102741	U79291	AW959829	Hs.83572	hypothetical protein MGC14433
45	101175	U82671_cds2	U82671	Hs.36980	melanoma antigen, family A, 2
	132164	U84573	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dioxygenase (lysine hydroxylase) 2
	102823	U90914	D85390	Hs.5057	carboxypeptidase D
	102826	U91316	NM_007274	Hs.8679	cytosolic acyl coenzyme A thioester hydrolase
	102831	U91932	AA262170	Hs.80917	adaptor-related protein complex 3, sigma 1 subunit
50	102846	U96131	BE264974	Hs.6566	thyroid hormone receptor Interactor 13
	129777	U97018	U97018	Hs.12451	echinoderm microtubule-associated protein-like
	134161	U97188	AA634543	Hs.79440	IGF-II mRNA-binding protein 3
	134854	V00503	J03464	Hs.179573	collagen, type I, alpha 2
	302363	X04327	AW163799	Hs.198365	2,3-bisphosphoglycerate mutase
55	133708	X06389	AI018666	Hs.75667	synaptophysin
	125701	X07496	T72104	Hs.93194	apolipoprotein A-I
	102915	X07820	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin 2)
	134656	X14787	AI750878	Hs.87409	thrombospondin 1
	413858	X15525_ma1	NM_001610	Hs.75589	acid phosphatase 2, lysosomal
60	102968	X16396	AU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase (NAD+ dependent), methylenetetrahydrofolate
					cyclohydrolase
	102971	X16609	X16609	Hs.183805	ankyrin 1, erythrocytic
	134037	X53586_ma1	AI808780	Hs.227730	Integrin, alpha 6
	103023	X53793	AW500470	Hs.117950	multifunctional polypeptide similar to SAICAR synthetase and AIR carboxylase
65	103037	X54936	BE018302	Hs.2894	placental growth factor, vascular endothelial growth factor-related protein
	130282	X55740	BE245380	Hs.153952	5' nucleotidase (CD73)
	134542	X57025	M14156	Hs.85112	insulin-like growth factor 1 (somatomedin C)
	128568	X60673_ma1	H12912	Hs.274691	adenylate kinase 3
	103093	X60708	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2)
70	133606	X62048	U10564	Hs.75188	wee1+ (S. pombe) homolog
	128063	X63097	X63094	Hs.283822	Rhesus blood group, D antigen
	424460	X63563	BE275979	Hs.296014	polymerase (RNA) II (DNA directed) polypeptide B (140kD)
	133227	X64037	AW977263	Hs.68257	general transcription factor IIF, polypeptide 1 (74kD subunit)
	103181	X69636	X69636	Hs.334731	Homo sapiens, clone IMAGE:3448306, mRNA, partial cds
75	103184	X69878	U43143	Hs.74049	fms-related tyrosine kinase 4
	103194	X70649	NM_004939	Hs.78580	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1

5	103208	X72841	AW411340	Hs.31314	retinoblastoma-binding protein 7
	129698	X74987	BE242144	Hs.12013	ATP-binding cassette, sub-family E (OABP), member 1
	131486	X83107	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	130729	X84194	AI963747	Hs.18573	acylphosphatase 1, erythrocyte (common) type
	103334	X85753	NM_001260	Hs.25283	cyclin-dependent kinase 8
	132645	X87670	AI654712	Hs.54424	hepatocyte nuclear factor 4, alpha
	135094	X89066	NM_003304	Hs.250687	transient receptor potential channel 1
	103352	X89398_cds2	H09366	Hs.78853	uracil-DNA glycosylase
10	103353	X89399	X89399	Hs.119274	RAS p21 protein activator (GTPase activating protein) 3 (ins(1,3,4,5)P4-binding protein)
	132173	X89426	X89426	Hs.41716	endothelial cell-specific molecule 1
	103371	X91247	X91247	Hs.13046	thioredoxin reductase 1
	131584	X91648	AA598509	Hs.29117	purine-rich element binding protein A
	103376	X92098	AL036166	Hs.323378	coated vesicle membrane protein
	103378	X92110	AL119690	Hs.153618	HCGVIII-1 protein
15	128510	X94703	X94703	Hs.296371	RAB28, member RAS oncogene family
	103410	X96506	AA158294	Hs.334879	DR1-associated protein 1 (negative cofactor 2 alpha)
	133490	X97230_f	AF022044	Hs.274601	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 1
	103438	X98263	AW175781	Hs.152720	M-phase phosphoprotein 6
	103440	X98296	X98296	Hs.77578	ubiquitin specific protease 9, X chromosome (Drosophila fat facets related)
20	103452	X99584	NM_006936	Hs.85119	SMT3 (suppressor of mit two 3, yeast) homolog 1
	133536	Y00264	W25797.comp	Hs.177486	amyloid beta (A4) precursor protein (protease nexin-II, Alzheimer disease)
	135185	Y07566	AW404908	Hs.96038	Ric (Drosophila)-like, expressed in many tissues
	118523	Y07759	Y07759	Hs.170157	myosin VA (heavy polypeptide 12, myosin)
	134662	Y07827	NM_007048	Hs.284283	butyrophilin, subfamily 3, member A1
25	132083	Y07867	BE386490	Hs.279863	Pirin
	103500	Y09443	AW408009	Hs.22580	alkylglycerone phosphate synthase
	134389	Y09658	Y09658	Hs.82577	spindlin-like
	132084	Y12394	NM_002267	Hs.3886	karyopherin alpha 3 (Importin alpha 4)
30	103540	Z11559	NM_002197	Hs.154721	aconitase 1, soluble
	133152	Z11695	Z11695	Hs.324473	mitogen-activated protein kinase 1
	103548	Z15005	Z15005	Hs.75573	centromere protein E (312kD)
	103612	Z46261	BE336654	Hs.70937	H3 histone family, member A
	129092	AA011243_s	D56365	Hs.63525	poly(rC)-binding protein 2
35	103692	AA018416	AW137912	Hs.227583	Homo sapiens chromosome X map Xp11.23 L-type calcium channel alpha-1 subunit
	(CACNA1F) gene, complete cds; HSP27 pseudogene, complete sequence; and JM1 protein, JM2 protein, and Hb2E genes, complete cds				
	103695	AA018758	AW207152	Hs.186600	ESTs
	129796	AA018804	BE218319	Hs.5807	GTPase Rab14
	132258	AA031993	AA306325	Hs.4311	SUMO-1 activating enzyme subunit 2
	132683	AA044217	BE264633	Hs.143638	WD repeat domain 4
40	131887	AA046548	W17064	Hs.332848	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1
	103723	AA057447_s	BE274312	Hs.214783	Homo sapiens cDNA FLJ14041 fis, clone HEMBA1005780
	453368	AA058376	W20296	Hs.288178	Homo sapiens cDNA FLJ11968 fis, clone HEMBB1001133
45	133260	AA083572	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone REC00917
	103765	AA085696	AA085696	Hs.169600	KIAA0826 protein
	103766	AA088744	AI920783	Hs.191435	ESTs
	103767	AA089688	BE244667	Hs.296155	CGI-100 protein
	132051	AA091284	AA393968	Hs.180145	HSPC030 protein
50	103773	AA092700	AI219323	Hs.101077	ESTs, Weakly similar to T22363 hypothetical protein F47G9.4 - Caenorhabditis elegans
	[C.elegans]				
	135289	AA092968	AW372569	Hs.9788	hypothetical protein MGC10924 similar to Nedd4 WW-binding protein 5
	132729	AA094800	AW970843	Hs.55682	eukaryotic translation initiation factor 3, subunit 7 (zeta, 66/67kD)
	103794	AA100219	AF244135	Hs.30670	hepatocellular carcinoma-associated antigen 66
55	131471	AA114885	AA164842	Hs.192619	KIAA1600 protein
	134319	AA129547	BE304999	Hs.75653	fumarate hydratase
	103807	AA133016	AW958264	Hs.103832	similar to yeast Upf3, variant B
	119159	AA149507	AF142419	Hs.15020	homolog of mouse quaking QKI (KH domain RNA binding protein)
	129863	AA151005	BE379765	Hs.129872	sperm associated antigen 9
60	103850	AA187101	AA187101	Hs.213194	hypothetical protein MGC10895
	103855	AA195179_s	W02363	Hs.302267	hypothetical protein FLJ10330
	322026	AA203138	AW024973	Hs.283675	NPD009 protein
	135300	AA203645	AA142922	Hs.278626	Arg/Abi-interacting protein ArgBP2
	103861	AA206236	AA206236	Hs.4944	hypothetical protein FLJ12783
65	130634	AA227621	AI769067	Hs.127824	ESTs, Weakly similar to T28770 hypothetical protein W03D2.1 - Caenorhabditis elegans
	[C.elegans]				
	447735	AA248283	AA775268	Hs.6127	Homo sapiens cDNA: FLJ23020 fis, clone LNG00943
	103909	AA249611	AA249611	Hs.47438	SH3 domain binding glutamic acid-rich protein
	131236	AA282640	AF043117	Hs.24594	ubiquitination factor E4B (homologous to yeast UFD2)
	134060	AA287199	D42039	Hs.78871	mesoderm development candidate 2
70	129013	AA313990	AA371156	Hs.107942	DKFZP564M112 protein
	129435	AA314256	AF151852	Hs.111449	CGI-94 protein
	103988	AA314389	AA314389	Hs.42500	ADP-ribosylation factor-like 5
	104000	AA324364	AI146527	Hs.80475	polymerase (RNA) II (DNA directed) polypeptide J (13.3kD)
	425284	AA329211_s	AF155568	Hs.155489	NS1-associated protein 1
75	128629	AA399187	AL096748	Hs.102708	DKFZP434A043 protein
	133281	AA421079	AK001601	Hs.69594	high-mobility group 20A

	104104	AA422029	AA422029	Hs.143640	ESTs, Weakly similar to hyperpolarization-activated cyclic nucleotide-gated channel HCN2
	[H.sapiens]				
	108154	AA425230	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain-binding protein
	132081	AA447052	AW954243	Hs.170218	KIAA0251 protein
5	135073	AA452000	W55958	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (from clone DKFZp586E1624)
	131387	AA456687	AI750575	Hs.173933	nuclear factor I/A
	129593	AA487015_s	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (from clone DKFZp586L0120)
	135266	AB002326	R41179	Hs.97353	KIAA0328 protein
	133505	C01527	AI630124	Hs.324504	Homo sapiens mRNA; cDNA DKFZp586J0720 (from clone DKFZp586J0720)
10	132064	C01714	AA121098	Hs.3838	serum-inducible kinase
	134393	C01811_J	W52642	Hs.8261	hypothetical protein FLJ22393
	131427	C02352_s	AF151879	Hs.28706	CGI-121 protein
	133435	C02375	AI929357	Hs.323966	Homo sapiens clone H63 unknown mRNA
	104282	C14448	C14448	Hs.332338	EST
15	134827	D16611_s	BE314037	Hs.89866	coproporphyrinogen oxidase (coproporphyrin, harderoporphyrin)
	130443	D25218	D25218	Hs.155650	KIAA0014 gene product
	131742	D31352	AA961420	Hs.31433	ESTs
	132837	D58024_s	AA370362	Hs.57958	EGF-TM7-latrophilin-related protein
	130377	D80897	NM_014909	Hs.155182	KIAA1036 protein
20	104334	D82614	D82614	Hs.78771	phosphoglycerate kinase 1
	134593	D87845	NM_000437	Hs.234392	platelet-activating factor acetylhydrolase 2 (40kD)
	134731	D89377_J	D89377	Hs.89404	msh (Drosophila) homeo box homolog 2
	129913	H06583	NM_001310	Hs.13313	cAMP responsive element binding protein-like 2
	131670	H40732	H03514	Hs.10130	ESTs
25	104394	H46617	AA129551	Hs.172129	Homo sapiens cDNA: FLJ21409 fls, clone COL03924
	104402	H56731	H56731	Hs.132956	ESTs
	129781	H75570	AA306090	Hs.124707	ESTs
	129077	H78886	N74724	Hs.108479	ESTs
	104417	H81241	AI819448	Hs.320861	Kruppel-like factor 8
30	134927	L36531	L36531	Hs.91296	integrin, alpha 8
	129280	M63154	M63154	Hs.110014	gastric intrinsic factor (vitamin B synthesis)
	134498	M63180	AW246273	Hs.84131	threonyl-tRNA synthetase
	104460	M61504	AW955705	Hs.82604	Homo sapiens, clone IMAGE:4299322, mRNA, partial cds
	104488	N56191	N56191	Hs.106511	protocadherin 17
35	131248	N78483	AI038989	Hs.332633	Bardet-Biedl syndrome 2
	129214	N79268	AL044335	Hs.109526	zinc finger protein 198
	130017	R14652	AK000096	Hs.143198	inhibitor of growth family, member 3
	104530	R20458	AK001676	Hs.12457	hypothetical protein FLJ10814
40	104534	R22303	R22303		gbz699.1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:130841 5', mRNA sequence.
	104544	R33779	AI091173	Hs.222362	ESTs, Weakly similar to p40 [H.sapiens]
	133328	R36553	AW452738	Hs.265327	hypothetical protein DKFZp761141
	104567	R64534	AA040620	Hs.5672	hypothetical protein AF140225
45	128562	R66475	AA923382	Hs.101490	ESTs
	129575	R70621	F08282	Hs.278428	progesterone induced protein
	130776	R79356	AF167706	Hs.19280	cysteine-rich motor neuron 1
	104599	R84933	AW815036	Hs.151251	ESTs
	104660	RC_AA007160	BE298665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (from clone DKFZp564D016)
50	104667	RC_AA007234_s	AI239923	Hs.30098	ESTs
	104718	RC_AA018409	AI143020	Hs.36250	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	104764	RC_AA025351	AI039243	Hs.278585	ESTs
	104786	RC_AA027168	AA027167	Hs.10031	KIAA0955 protein
	104787	RC_AA027317	AA027317		gbz697d11.1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:366933 3'
55	134079	RC_AA029423	AK001751	Hs.171835	hypothetical protein FLJ10889
	104804	RC_AA031357	AI858702	Hs.31803	ESTs, Weakly similar to N-WASP [H.sapiens]
	104865	RC_AA045136	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc finger protein)
	130828	RC_AA053400	AW631469	Hs.203213	ESTs
60	104907	RC_AA055829	AA055829	Hs.196701	ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
	WARNING ENTRY [H.sapiens]				
	104943	RC_AA065217	AF072873	Hs.114218	frizzled (Drosophila) homolog 6
	105013	RC_AA116054	H63789	Hs.296288	ESTs, Weakly similar to KIAA0838 protein [H.sapiens]
	105024	RC_AA126311	AA126311	Hs.9879	ESTs
65	132592	RC_AA129390	AW803564	Hs.288850	Homo sapiens cDNA: FLJ22528 fls, clone HRC12825
	105038	RC_AA130273	AW503733	Hs.9414	KIAA1468 protein
	105077	RC_AA142919	W55948	Hs.234863	Homo sapiens cDNA FLJ12082 fls, clone HEMBB1002492
	105096	RC_AA150205	AL042506	Hs.21599	Kruppel-like factor 7 (ubiquitous)
	129215	RC_AA176887	AB040930	Hs.126085	KIAA1497 protein
	105169	RC_AA180321	BE245294	Hs.180789	S164 protein
70	132796	RC_AA180487	NM_006283	Hs.173159	transforming, acidic coiled-coil containing protein 1
	130401	RC_AA187634	BE396283	Hs.173987	eukaryotic translation initiation factor 3, subunit 1 (alpha, 35kD)
	105200	RC_AA195399	AA328102	Hs.24641	cytoskeleton associated protein 2
	130114	RC_AA234717	AA233393	Hs.14992	hypothetical protein FLJ11151
	105330	RC_AA234743	AW338625	Hs.22120	ESTs
75	105337	RC_AA234957	AI468789	Hs.23200	myotubularin related protein 1
	129385	RC_AA235604	AA172106	Hs.110950	Rag C protein

	105376	RC_AA236559	AW994032	Hs.8768	hypothetical protein FLJ10849
	105397	RC_AA242868	AA814807	Hs.7395	hypothetical protein FLJ23182
	131862	RC_AA251776	AK000046	Hs.267448	hypothetical protein FLJ20039
	131991	RC_AA251909	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1 (yeast homolog), beta
5	128658	RC_AA252872_s	BE397354	Hs.324830	diphtheria toxin resistance protein required for diphthamide biosynthesis (Saccharomyces)-like 2
	105489	RC_AA256157	AA256157	Hs.24115	Homo sapiens cDNA FLJ14178 fis, clone NT2RP2003339
	105508	RC_AA256680	AA173942	Hs.326416	Homo sapiens mRNA; cDNA DKFZp564H1916 (from clone DKFZp564H1916)
	105539	RC_AA258873	AB040884	Hs.109694	KIAA1451 protein
	135172	RC_AA262727	AB028956	Hs.12144	KIAA1033 protein
10	131569	RC_AA281451	AL389951	Hs.271623	nucleoporin 50kD
	132542	RC_AA281545	AL137751	Hs.263671	Homo sapiens mRNA; cDNA DKFZp434I0812 (from clone DKFZp434I0812); partial cds
	105643	RC_AA282069	BE621719	Hs.173802	KIAA0603 gene product
	105659	RC_AA283044	AA283044	Hs.25625	hypothetical protein FLJ11323
	105666	RC_AA283930	AA426234	Hs.34906	ESTs, Weakly similar to T17210 hypothetical protein DKFZp434N041.1 [H.sapiens]
15	105674	RC_AA284755	AI609530	Hs.279789	histone deacetylase 3
	105709	RC_AA291268	AI928962	Hs.26761	DKFZP586L0724 protein
	105722	RC_AA291927	AI922821	Hs.32433	ESTs
	105765	RC_AA343514	AA299688	Hs.24183	ESTs
	115951	RC_AA398109	BE546245	Hs.301048	sec13-like protein
20	105962	RC_AA405737	AW880358	Hs.339808	hypothetical protein FLJ10120
	105985	RC_AA406610	AA406610		gbzv15b10.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:753691 3' similar to
	106008	RC_AA411465	AB033888	Hs.8619	SRY (sex determining region Y)-box 18
	131216	RC_AA416886	AI815486	Hs.243901	Homo sapiens cDNA FLJ20738 fis, clone HEP08257
25	134222	RC_AA424013	AW855861	Hs.8025	Homo sapiens clone Z3767 and Z3782 mRNA sequences
	113689	RC_AA424148	AB037850	Hs.16621	DKFZP434I116 protein
	106141	RC_AA424558	AF031463	Hs.9302	phosphatase-like
	130839	RC_AA424961_s	AB011169	Hs.20141	similar to S. cerevisiae SSM4
	106157	RC_AA425367	W37943	Hs.34892	KIAA1323 protein
30	130777	RC_AA425921	AW135049	Hs.285418	Homo sapiens cDNA FLJ10643 fis, clone NT2RP2005753, highly similar to Homo sapiens I-1
	130561	RC_AA426220	AB011095	Hs.16032	KIAA0523 protein
	106196	RC_AA427735	AA525993	Hs.173699	ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
	131878	RC_AA430673	AA083764	Hs.6101	hypothetical protein MGC3178
35	133200	RC_AA432248	AB037715	Hs.183639	hypothetical protein FLJ10210
	106302	RC_AA435896	AA398859	Hs.18397	hypothetical protein FLJ23221
	106328	RC_AA436705	AL079559	Hs.28020	KIAA0766 gene product
	450534	RC_AA446561	AI570189	Hs.25132	KIAA0470 gene product
40	106423	RC_AA448238	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15
	133442	RC_AA448688	AL137663	Hs.7378	Homo sapiens mRNA; cDNA DKFZp434G227 (from clone DKFZp434G227)
	439608	RC_AA449756	AW864696	Hs.301732	hypothetical protein MGC5306
	106477	RC_AA450303	R23324	Hs.41693	DnaJ (Hsp40) homolog, subfamily B, member 4
	106503	RC_AA452411	AB033042	Hs.29679	cofactor required for Sp1 transcriptional activation, subunit 3 (130kD)
45	446999	RC_AA454566	AA151520	Hs.334822	hypothetical protein MGC4485
	106543	RC_AA454667	AA676939	Hs.69285	neurotrophin 1
	130010	RC_AA456437	AA301116	Hs.142638	nucleolar phosphoprotein Nopp34
	106589	RC_AA456646	AK000933	Hs.28661	Homo sapiens cDNA FLJ10071 fis, clone HEMBA1001702
	106593	RC_AA456826	AW296451	Hs.24605	ESTs
50	106596	RC_AA456981	AA452379	Hs.293552	ESTs, Moderately similar to ALU7_HUMAN ALU SUBFAMILY SQ SEQUENCE
	134655	RC_AA458959	AF265208	Hs.123090	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily f, member 1
	106636	RC_AA459950	AW958037	Hs.286	ribosomal protein L4
55	106654	RC_AA460449	AW075485	Hs.286049	phosphoserine aminotransferase
	131353	RC_AA463910	AW754182		gbzRC2-CT0321-131199-011-c01 CT0321 Homo sapiens cDNA, mRNA sequence
	106707	RC_AA464603	AK000566	Hs.98135	hypothetical protein FLJ20559
	131710	RC_AA464606	NM_015368	Hs.30985	pannexin 1
	106717	RC_AA465093	AA600357	Hs.239489	TIA1 cytotoxic granule-associated RNA-binding protein
60	131775	RC_AA465692	AB014548	Hs.31921	KIAA0648 protein
	106747	RC_AA476473	NM_007118	Hs.171957	triple functional domain (PTPRF interacting)
	106773	RC_AA478109	AA478109	Hs.188833	ESTs
	106781	RC_AA478474	AA330310	Hs.24181	ESTs
	106817	RC_AA480889	D61216	Hs.18672	ESTs
65	106846	RC_AA485223	AB037744	Hs.34892	KIAA1323 protein
	106848	RC_AA485254	AA449014	Hs.121025	chromosome 11 open reading frame 5
	106856	RC_AA486183	W58353	Hs.285123	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 2005779
	418699	RC_AA486936	BE539639	Hs.173030	ESTs, Weakly similar to ALU8_HUMAN ALU SUBFAMILY SX SEQUENCE CONTAMINATION
	107001	RC_AA498589	AI926520	Hs.31016	putative DNA binding protein
70	130638	RC_AA498831_f	AW021278	Hs.17121	ESTs
	107054	RC_AA600150	AI076459	Hs.15978	KIAA1272 protein
	107059	RC_AA608545	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E. coli RecA homolog)
	107080	RC_AA609210	AL122043	Hs.19221	hypothetical protein DKFZp586G1424
75	107115	RC_AA610108	BE379623	Hs.27693	peptidylprolyl isomerase (cyclophilin)-like 1
	107130	RC_AA620582	AB033106	Hs.12913	KIAA1280 protein

	107156	RC_AA621239	AA137043	Hs.9663	programmed cell death 6-interacting protein
	107174	RC_AA621714	BE122762	Hs.25338	ESTs
	130621	RC_AA621718	AW513087	Hs.16803	LUC7 (S. cerevisiae)-like
	107190	RC_D19673	AA836401	Hs.5103	ESTs
5	132626	RC_D25755_s	AW504732	Hs.21275	hypothetical protein FLJ11011
	107217	RC_D51095	AL080235	Hs.35861	DKFZP586E1621 protein
	131610	RC_D60272_J	AA357879	Hs.29423	scavenger receptor with C-type lectin
	129604	T08879	AF088886	Hs.11590	cathepsin F
10	107295	T34527	AA186629	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1 (GalNAc-T1)
	107299	T40327_s	BE277457	Hs.30661	hypothetical protein MGC4606
	107315	T62771_s	AA316241	Hs.90691	nucleophosmin/nucleoplasm 3
	107316	T63174_s	T63174	Hs.193700	Homo sapiens mRNA; cDNA DKFZp586i0324 (from clone DKFZp586i0324)
	107328	T83444	AW959891	Hs.76591	KIAA0887 protein
15	107334	T93641	T93597	Hs.167429	ESTs
	134715	U48263	U48263	Hs.89040	prepronodectin
	128636	U49065	U49065	Hs.102865	interleukin 1 receptor-like 2
	129938	U79300	AW003668	Hs.135587	Human clone Z3629 mRNA sequence
	107375	U88573	BE011845	Hs.251064	high-mobility group (nonhistone chromosomal) protein 14
20	130074	U93867	AL038596	Hs.250745	polymerase (RNA) III (DNA directed) (62kD)
	107387	W01094	D86983	Hs.118893	Melanoma associated gene
	132036	W01568	AL157433	Hs.37708	hypothetical protein DKFZp434E2220
	107426	W26853	W26853	Hs.291003	hypothetical protein MGC4707
	113857	W27179	AW243158	Hs.5297	DKFZP564A2416 protein
25	135388	W27965	W27965	Hs.99865	epimorphin
	130419	W36280_s	AF037448	Hs.155489	NS1-associated protein 1
	107469	W47063	W47063	Hs.94668	ESTs
	132616	W79060	BE262677	Hs.283558	hypothetical protein PRO1855
	107506	W88550	AB028981	Hs.8021	KIAA1058 protein
30	132358	X60486	NM_003542	Hs.46423	H4 histone family, member G
	107522	X78931_s	X78931	Hs.99971	zinc finger protein 272
	125827	Z14077_s	NM_003403	Hs.97496	YY1 transcription factor
	107582	RC_AA002147	AA002147	Hs.59952	EST
	107609	RC_AA004711	R75654	Hs.164797	hypothetical protein FLJ13693
35	107661	RC_AA010383	AA010383	Hs.60389	ESTs
	107714	RC_AA015761	AA015761	Hs.60642	ESTs
	107775	RC_AA018772	AW008846	Hs.60857	ESTs
	107832	RC_AA021473_J	AA021473		gb:ze66c11.s1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:363956 3', mRNA sequence.
40	107859	RC_AA024835	AW732573	Hs.47584	potassium voltage-gated channel, delayed-rectifier, subfamily S, member 3
	124337	RC_AA025858	N23541	Hs.281561	Homo sapiens cDNA: FLJ23582 tis, clone LNG13759
	107914	RC_AA027229	AA027229	Hs.61329	ESTs, Weakly similar to T16370 hypothetical protein F45E12.5 - Caenorhabditis elegans
		[C.elegans]			
	107935	RC_AA029428	AA029428	Hs.61555	ESTs
45	116262	RC_AA035143	A1936442	Hs.59838	hypothetical protein FLJ10808
	131461	RC_AA035237	AA992841	Hs.27263	KIAA1458 protein
	108007	RC_AA039347	AA039347	Hs.61916	EST
	108029	RC_AA040740	AA040740	Hs.62007	ESTs
	108040	RC_AA041551	AL121031	Hs.159971	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1
50	108084	RC_AA045513	AA058944	Hs.116602	Homo sapiens, clone IMAGE:4154008, mRNA, partial cds
	108088	RC_AA045745	AA045745	Hs.62886	ESTs
	108168	RC_AA055348	A1453137	Hs.63176	ESTs
	130719	RC_AA056582_s	AA679262	Hs.14235	hypothetical protein FLJ20008; KIAA1839 protein
55	108189	RC_AA056697	AW376061	Hs.63335	ESTs, Moderately similar to A46010 X-linked retinopathy protein [H.sapiens]
	108190	RC_AA056746	AA056746	Hs.63338	EST
	108203	RC_AA057678	AW847814	Hs.289005	Homo sapiens cDNA: FLJ21532 tis, clone COL06049
	108216	RC_AA058681	AA524743	Hs.44883	ESTs
	108217	RC_AA058686	AA058686	Hs.62588	ESTs
60	108245	RC_AA062840	BE410285	Hs.89545	proteasome (prosome, macropain) subunit, beta type, 4
	108277	RC_AA064859	AA064859		gb:zm50i03.s1 Stratagene fibroblast (937212) Homo sapiens cDNA clone IMAGE:529085 3', mRNA
	108280	RC_AA065069	AA065069		gb:zm12e11.s1 Stratagene pancreas (937208) Homo sapiens cDNA clone 3', mRNA sequence
	108309	RC_AA069923	AA069818		gb:zm67e03.r1 Stratagene neuroepithelium (937231) Homo sapiens cDNA clone 5' similar to
65	133739	RC_AA070799_s	BE536554	Hs.278270	inactive progesterone receptor, 23 kD
	108340	RC_AA070815	AA069820	Hs.180909	peroxiredoxin 1
	108403	RC_AA075374	AA075374		gb:zm87e01.s1 Stratagene ovarian cancer (937219) Homo sapiens cDNA clone IMAGE:544872
		3', mRNA sequence.			
	108427	RC_AA076382	AA076382		gb:zm91g08.s1 Stratagene ovarian cancer (937219) Homo sapiens cDNA clone IMAGE:545342
		3', mRNA sequence.			
70	108435	RC_AA078787	T82427	Hs.194101	Homo sapiens cDNA: FLJ20869 tis, clone ADKA02377
	108439	RC_AA078986	AA078986		gb:zm92h01.s1 Stratagene ovarian cancer (937219) Homo sapiens cDNA clone IMAGE:545425
		3', mRNA sequence.			
	108465	RC_AA079393	AA079393	Hs.3462	cytochrome c oxidase subunit VIc
75	108469	RC_AA079487	AA079487		gb:zm97f08.s1 Stratagene colon HT29 (937221) Homo sapiens cDNA clone 3', mRNA sequence

	108500	RC_AA083207	AA083207	Hs.68270	EST
	108501	RC_AA083256	AA083256		gbzm08g12.s1 Stratagene hNT neuron (937233) Homo sapiens cDNA clone 3' similar to
	gb:M33308				
5	108533	RC_AA084415	AA084415		gbzm06g09.s1 Stratagene hNT neuron (937233) Homo sapiens cDNA clone IMAGE:546688 3',
	mRNA				
	108562	RC_AA085274	AA100796		gbzm26c06.s1 Stratagene pancreas (937208) Homo sapiens cDNA clone 3' similar to
	gb:X15341				
	108589	RC_AA088678	AI732404	Hs.68846	ESTs
	130890	RC_AA100925	AI907537	Hs.76698	stress-associated endoplasmic reticulum protein 1; ribosome associated membrane protein 4
10	134585	RC_AA101255	D14041	Hs.278573	H-2K binding factor-2
	130385	RC_AA126474	AW067800	Hs.155223	stannocalcin 2
	108749	RC_AA127017	AA127017	Hs.71052	ESTs
	108807	RC_AA129968	AI552236	Hs.49376	hypothetical protein FLJ20644
	108808	RC_AA130240	AA045088	Hs.62738	ESTs
15	108833	RC_AA131866	AF188527	Hs.61661	ESTs, Weakly similar to AF174605 1 F-box protein Fbx25 [H.sapiens]
	107290	RC_AA132039	W27740	Hs.323780	ESTs
	108846	RC_AA132983	AL117452	Hs.44155	DKFZP586G1517 protein
	108857	RC_AA133250	AK001468	Hs.62180	anillin (Drosophila Scaps homolog), actin binding protein
	131474	RC_AA133583_s	L46353	Hs.2726	high-mobility group (nonhistone chromosomal) protein isoform I-C
20	108894	RC_AA135941	AK001431	Hs.5105	hypothetical protein FLJ10569
	108941	RC_AA148650	AA148650		gbzo09e06.s1 Stratagene neuroepithelium NT2RAM1 937234 Homo sapiens cDNA clone
	IMAGE:567202 3',				
	108968	RC_AA151110	AI304870	Hs.188680	ESTs
	108996	RC_AA155754	AW995610	Hs.332436	EST
25	109001	RC_AA156125	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to hedgehog-interacting protein
	131183	RC_AA156289	AI611807	Hs.285107	hypothetical protein FLJ13397
	109019	RC_AA156997	AA156755	Hs.72150	ESTs
	109022	RC_AA157291	AA157291	Hs.21479	ubiquitin 1
	109023	RC_AA157293	AA157293	Hs.72168	ESTs
30	109068	RC_AA164293_f	AA164293	Hs.72545	ESTs
	109072	RC_AA164676	AI732585	Hs.22394	hypothetical protein FLJ10893
	129021	RC_AA167375	AL044675	Hs.173081	KIAA0530 protein
	130346	RC_AA167550	H05769	Hs.188757	Homo sapiens, clone MGC:5564, mRNA, complete cds
	109146	RC_AA176589	AA176589	Hs.142078	EST
35	109172	RC_AA180448	AA180448	Hs.144300	EST
	131080	RC_AA187144_s	NM_001955	Hs.2271	endothelin 1
	129208	RC_AA189170_f	AI587376	Hs.109441	MSTP033 protein
	109222	RC_AA192757	AA192833	Hs.333512	similar to rat myomegalin
	109300	RC_AA205650	AA418276	Hs.170142	ESTs
40	109481	RC_AA233342	AA878923	Hs.289069	hypothetical protein FLJ21016
	109485	RC_AA233472	BE619092	Hs.28465	Homo sapiens cDNA: FLJ21869 tis, clone HEP02442
	109516	RC_AA234110	AI471639	Hs.71913	ESTs
	109537	RC_D80981	AI858695	Hs.34898	ESTs
	109556	RC_F01660	AI925294	Hs.87385	ESTs
45	109577	RC_F02206	F02206	Hs.296639	Homo sapiens potassium channel subunit (HERG-3) mRNA, complete cds
	109578	RC_F02208	F02208	Hs.27214	ESTs
	109595	RC_F02544	AA078629	Hs.27301	ESTs
	109625	RC_F03918	H29490	Hs.22697	ESTs
	131983	RC_F04258_s	AF118665	Hs.184011	pyrophosphatase (inorganic)
50	109648	RC_F04600	H17800	Hs.7154	ESTs
	109671	RC_F08998	R59210	Hs.26634	ESTs
	109699	RC_F09605	H18013	Hs.167483	ESTs
	109820	RC_F11115	AW016809	Hs.323795	ESTs
	109933	RC_H06371	R52417	Hs.20945	Homo sapiens clone 24993 mRNA sequence
55	110014	RC_H10995	AL109666	Hs.7242	Homo sapiens mRNA full length insert cDNA clone EUROMIMAGE 35907
	110039	RC_H11938	H11938	Hs.21907	histone acetyltransferase
	110099	RC_H16568	R44557	Hs.23748	ESTs
	110107	RC_H16772	AW151660	Hs.31444	ESTs
	110155	RC_H18951	AI559826	Hs.93522	Homo sapiens mRNA for KIAA1647 protein, partial cds
60	110197	RC_H20859	AW090386	Hs.112278	arrestin, beta 1
	110223	RC_H23747	H19836	Hs.31697	ESTs
	110306	RC_H38087	H38087	Hs.105509	CTL2 gene
	110335	RC_H40331	H65490	Hs.18845	ESTs
	110342	RC_H40567	H40561	Hs.33008	ESTs
65	110395	RC_H46966	AA025116	Hs.33333	ESTs
	110511	RC_H56640_J	H56640	Hs.221460	ESTs
	110523	RC_H57154	AI040384	Hs.19102	ESTs, Weakly similar to organic anion transporter 1 [H.sapiens]
	110715	RC_H96712	H96712	Hs.269029	ESTs
	110754	RC_N20814	AW302200	Hs.6336	KIAA0672 gene product
70	130132	RC_N25249	U55936	Hs.184376	synapsoosomal-associated protein, 23kD
	131135	RC_N27100	NM_016569	Hs.267182	TBX3-iso protein
	134263	RC_N39616	AW973443	Hs.8086	RNA (guanine-7-) methyltransferase
	110938	RC_N48982	N48982	Hs.38034	Homo sapiens cDNA FLJ12924 tis, clone NT2RP2004709
	110983	RC_N51957	NM_015367	Hs.10267	MIL1 protein
75	115062	RC_N52271	AA253314	Hs.154103	LIM protein (similar to rat protein kinase C-binding enigma)
	111081	RC_N59435	AI146349	Hs.271614	CGI-112 protein

5	111128 RC_N64139	AW505364	Hs.19074	LATS (large tumor suppressor, Drosophila) homolog 2
	135244 RC_N66981	A1834273	Hs.9711	novel protein
	111216 RC_N68640	AW139408	Hs.152940	ESTs
	437582 RC_N69352	AB001636	Hs.5683	DEAD/HD (Asp-Glu-Ala-Asp/His) box polypeptide 15
	131002 RC_N65226	AL050295	Hs.22039	KIAA0758 protein
	111399 RC_R00138	AW270776	Hs.18857	ESTs
	111514 RC_R07998	R07998		gb:ylf16g11.1.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:127076 3'
	similar to			
10	130182 RC_R08929	BE267033	Hs.192853	ubiquitin-conjugating enzyme E2G 2 (homologous to yeast UBC7)
	111574 RC_R10307	AJ024145	Hs.188526	ESTs
	111804 RC_R33354	AA482478	Hs.181785	ESTs
	111831 RC_R36083	R36095	Hs.268695	ESTs
	129675 RC_R37938_f	NM_015556	Hs.172180	KIAA0440 protein
	111904 RC_R39330	Z41572		gb:HSCZYB122 normalized infant brain cDNA Homo sapiens cDNA clone c-zyb12, mRNA
15	sequence			
	133868 RC_R40816_s	AB012193	Hs.183874	cutin 4A
	112033 RC_R43162_s	R49031	Hs.22627	ESTs
	130987 RC_R45698	BE613269	Hs.21893	hypothetical protein DKFZp761N0624
20	112300 RC_R54554	H24334	Hs.26125	ESTs
	112513 RC_R68425	R68425	Hs.13809	hypothetical protein FLJ10648
	112514 RC_R68568	R68568	Hs.183373	src homology 3 domain-containing protein HIP-55
	112522 RC_R68763	R68857	Hs.265499	ESTs
	112540 RC_R70467	R69751		gb:Y40a10.s1 Soares placenta Nb2HP Homo sapiens cDNA clone 3', mRNA sequence
25	130346 RC_R73565	H05769	Hs.188757	Homo sapiens, clone MGC:5564, mRNA, complete cds
	129534 RC_R73640	AK002126	Hs.11260	hypothetical protein FLJ11264
	112597 RC_R78376	R78376	Hs.29733	EST
	112732 RC_R92453	R92453	Hs.34590	ESTs
	131458 RC_T03865	BE297567	Hs.27047	hypothetical protein FLJ20392
	112888 RC_T03872	AW195317	Hs.107716	hypothetical protein FLJ22344
30	131863 RC_T10072	AI656378	Hs.33461	ESTs
	112911 RC_T10080	AW732747	Hs.13493	like mouse brain protein E46
	132215 RC_T10132	AL035703	Hs.4236	KIAA0478 gene product
	112931 RC_T15343	T02968	Hs.167428	ESTs
35	112984 RC_T23457	T16971	Hs.289014	ESTs, Weakly similar to A43932 mucin 2 precursor, intestinal [H.sapiens]
	112998 RC_T23555	H11257	Hs.22968	Homo sapiens clone IMAGE:451939, mRNA sequence
	133376 RC_T23870	BE618768	Hs.7232	acetyl-Coenzyme A carboxylase alpha
	113026 RC_T23948	AA376654	Hs.183684	eukaryotic translation initiation factor 4 gamma, 2
	113070 RC_T33464	AB032977	Hs.6298	KIAA1151 protein
40	128970 RC_T34413	A1375672	Hs.165028	ESTs
	113074 RC_T34611	AK001335	Hs.31137	protein tyrosine phosphatase, receptor type, E
	113095 RC_T40920	AA828380	Hs.126733	ESTs
	113179 RC_T55182	BE622021	Hs.152571	ESTs, Highly similar to IGF-II mRNA-binding protein 2 [H.sapiens]
	113337 RC_T77453	T77453	Hs.302234	ESTs
45	113421 RC_T84039	AJ769400	Hs.189729	ESTs
	113454 RC_T86458	AJ022166	Hs.16188	ESTs
	113481 RC_T87693	T87693	Hs.204327	EST
	131441 RC_T89350_s	AA302862	Hs.90063	neurocalcin delta
50	113557 RC_T90845	H66470	Hs.16004	ESTs
	113559 RC_T90887	T79763	Hs.14514	ESTs
	113589 RC_T91863	AJ078554	Hs.15682	ESTs
	113591 RC_T91881	T91881	Hs.200597	KIAA0563 gene product
	113619 RC_T93783_s	R08665	Hs.17244	hypothetical protein FLJ13605
	113683 RC_T96887	AB035335	Hs.144519	T-cell leukemia/lymphoma 6
55	113692 RC_T96944	AL360143	Hs.17936	DKFZP434H132 protein
	113702 RC_T97307	T97307		gb:ye53h05.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:121497 3', mRNA
	113717 RC_T97764	T99513	Hs.187447	ESTs
	113824 RC_W48817	AI631964	Hs.34447	ESTs
60	113840 RC_W58343	R72137	Hs.7949	DKFZP566B2420 protein
	113844 RC_W59949	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HEMBB1001294, highly similar to GTP-BINDING
	PROTEIN TC10			
	113902 RC_W74644	AA340111	Hs.100009	acyl-Coenzyme A oxidase 1, palmitoyl
	113904 RC_W74761	AF125044	Hs.19196	ubiquitin-conjugating enzyme HBUCE1
65	113905 RC_W74802	R81733	Hs.33106	ESTs
	113931 RC_W81205	BE255499	Hs.3496	hypothetical protein MGC15749
	113932 RC_W81237	AA256444	Hs.126485	hypothetical protein FLJ12604; KIAA1692 protein
	131965 RC_W90146_f	W79283	Hs.35962	ESTs
	114036 RC_W92798	W92798	Hs.269181	ESTs
70	114106 RC_Z38412	AW602528		gb:RC5-BT0562-260100-011-A02 BT0562 Homo sapiens cDNA, mRNA sequence
	133593 RC_Z38709	AJ416988	Hs.238272	inositol 1,4,5-triphosphate receptor, type 2
	114161 RC_Z38904	BE548222	Hs.299883	hypothetical protein FLJ23399
	424949 RC_Z39103	AF052212	Hs.153934	core-binding factor, runt domain, alpha subunit 2; translocated to, 2
	129059 RC_Z39930_f	AW069534	Hs.279583	CGI-81 protein
75	128937 RC_Z39939	AA251380	Hs.10726	ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
	WARNING			
	130983 RC_Z40012_f	AJ479813	Hs.278411	NCK-associated protein 1

	114277	RC_Z40377_s	A1052229	Hs.25373	ESTs, Weakly similar to T20410 hypothetical protein E02A10.2 - <i>Caenorhabditis elegans</i>
		[C.elegans]			
	114304	RC_Z40820	A1934204	Hs.16129	ESTs
	114364	RC_Z41680	AL117427	Hs.172778	Homo sapiens mRNA; cDNA DKFZp566P013 (from clone DKFZp566P013)
5	132900	RC_AA005112	AA777749	Hs.5978	LIM domain only 7
	129034	RC_AA005432	AA481157	Hs.108110	DKFZP547E2110 protein
	131881	RC_AA010163	AW361018	Hs.3383	upstream regulatory element binding protein 1
	452461	RC_AA026356	N78223	Hs.108106	transcription factor
	114465	RC_AA026901	BE621056	Hs.131731	hypothetical protein FLJ11099
10	131376	RC_AA036867	AK001644	Hs.26156	hypothetical protein FLJ10782
	101567	RC_AA044644	M33552	Hs.56729	lysosomal
	431555	RC_AA046426	AI815470	Hs.260024	Cdc42 effector protein 3
	132944	RC_AA054515	T98641	Hs.6127	Homo sapiens cDNA: FLJ23020 fis, clone LNG00943
	114618	RC_AA084162	AW979261	Hs.291993	ESTs
15	130274	RC_AA085749	AA128376	Hs.153884	ATP binding protein associated with cell differentiation
	110330	RC_AA088874	AI288665	Hs.16621	DKFZP434I116 protein
	114648	RC_AA101056	AA101056		gbzr25b03.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone
	IMAGE:548429.3				
	114658	RC_AA102746	AA102383	Hs.249190	tumor necrosis factor receptor superfamily, member 10a
20	132456	RC_AA114250_s	AB011084	Hs.48924	KIAA0512 gene product; ALEX2
	131319	RC_AA126561_s	NM_003155	Hs.25590	stannocalcin 1
	132225	RC_AA128980_s	AA128980		gbzr25b03.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone
	IMAGE:567164.3				
	132669	RC_AA129757	W38586	Hs.293981	guanine nucleotide binding protein (G protein), gamma 3, linked
25	114709	RC_AA129921	AA397851	Hs.301959	proline synthetase co-transcribed (bacterial homolog)
	131973	RC_AA133331	AB018284	Hs.158688	KIAA0741 gene product
	114750	RC_AA135958	AA887211	Hs.129467	ESTs
	115714	RC_AA136524_s	T19228	Hs.172572	hypothetical protein FLJ20093
	114763	RC_AA147044	AA810755	Hs.88977	hypothetical protein dJ511E16.2
30	114767	RC_AA148885	AI859865	Hs.154443	minichromosome maintenance deficient (S. cerevisiae) 4
	114774	RC_AA150043	AV656017	Hs.184325	CGI-76 protein
	129388	RC_AA151621	AA662477	Hs.110964	hypothetical protein FLJ23471
	129183	RC_AA155743	BE561824	Hs.273369	uncharacterized hematopoietic stem/progenitor cells protein MDS027
	128869	RC_AA156335	AA768242	Hs.80618	hypothetical protein
35	130207	RC_AA156336	AF044209	Hs.144904	nuclear receptor co-repressor 1
	114788	RC_AA159181	AA159181	Hs.54900	serologically defined colon cancer antigen 1
	114800	RC_AA159825	Z19448	Hs.131887	ESTs, Weakly similar to T24396 hypothetical protein T03F6.2 - <i>Caenorhabditis elegans</i>
		[C.elegans]			
40	114828	RC_AA234185	AA252937	Hs.283522	Homo sapiens mRNA; cDNA DKFZp434J1912 (from clone DKFZp434J1912)
	114846	RC_AA234929	BE018682	Hs.166196	ATPase, Class I, type 8B, member 1
	114848	RC_AA234935	BE614347	Hs.169615	hypothetical protein FLJ20989
	114902	RC_AA236359	AW275480	Hs.39504	hypothetical protein MGC4308
	132271	RC_AA236466	AB030034	Hs.115175	sterile-alpha motif and leucine zipper containing kinase AZK
45	114907	RC_AA236535	N29390	Hs.13804	hypothetical protein dJ462O23.2
	135159	RC_AA236935_s	U43374	Hs.95631	Human normal keratinocyte mRNA
	132204	RC_AA236942	AA235827	Hs.42265	ESTs
	114928	RC_AA237018	AA237018	Hs.94869	ESTs
	132481	RC_AA237025	W93378	Hs.49614	ESTs
	114932	RC_AA242751	AA971436	Hs.16218	KIAA0903 protein
50	314162	RC_AA242760	BE041820	Hs.38516	Homo sapiens, clone MGC:15887, mRNA, complete cds
	131006	RC_AA242763	AF064104	Hs.22116	CDC14 (cell division cycle 14, S. cerevisiae) homolog B
	114935	RC_AA242809	H23329	Hs.290880	ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
		WARNING			
55	132454	RC_AA243133	BE296227	Hs.250822	serine/threonine kinase 15
	437754	RC_AA243495	R60366	Hs.5822	Homo sapiens cDNA: FLJ22120 fis, clone HEP18874
	114957	RC_AA243706	AW170425	Hs.87680	ESTs
	114974	RC_AA250848	AW966931	Hs.179662	nucleosome assembly protein 1-like 1
	114977	RC_AA250868	AW266978	Hs.87767	ESTs
60	114995	RC_AA251152	AA769266	Hs.193657	ESTs
	115005	RC_AA251544_s	AI760825	Hs.111339	ESTs
	417177	RC_AA251792	NM_004458	Hs.81452	fatty-acid-Coenzyme A ligase, long-chain 4
	131889	RC_AA252063	NM_002589	Hs.34073	BH-protocadherin (brain-heart)
	115026	RC_AA252144	AA251972	Hs.188718	ESTs
	115045	RC_AA252524	AW014549	Hs.58373	ESTs
65	115068	RC_AA253461	AW512260	Hs.87767	ESTs
	133138	RC_AA255522	AV657594	Hs.181161	Homo sapiens cDNA FLJ14643 fis, clone NT2RP2001597, weakly similar to RYANODINE
		RECEPTOR,			
	115114	RC_AA256468	AA527548	Hs.7527	small fragment nuclease
70	129584	RC_AA256528	AV656017	Hs.184325	CGI-76 protein
	115137	RC_AA257976	AW968304	Hs.56156	ESTs
	134312	RC_AA258296	AB011151	Hs.334659	hypothetical protein MGC14139
	115166	RC_AA258409	AF095727	Hs.287832	myelin protein zero-like 1
	115167	RC_AA258421	AA749209	Hs.43728	hypothetical protein
	129807	RC_AA262077	Y11192	Hs.5299	aldehyde dehydrogenase 5 family, member A1 (succinate-semialdehyde dehydrogenase)
75	115239	RC_AA278650	BE251328	Hs.73291	hypothetical protein FLJ10881
	115243	RC_AA278768	AA806600	Hs.116665	KIAA1842 protein

	100850	RC_AA279687_s	AA836472	Hs.297939	cathepsin B
	126884	RC_AA280791	U49436	Hs.286236	KIAA1856 protein
	115322	RC_AA280819	L08895	Hs.78995	MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C)
	133626	RC_AA280828	AW836130	Hs.75277	hypothetical protein FLJ13910
5	115372	RC_AA282195	AW014385	Hs.88678	ESTs, Weakly similar to Unknown [H.sapiens]
	132825	RC_AA283127_s	U82671	Hs.57698	Empirically selected from AFFX single probeset
	130269	RC_AA284694	F05422	Hs.168352	nucleoporin-like protein 1
	129192	RC_AA291137	AA286914	Hs.183299	ESTs
	452598	RC_AA291708	AI831594	Hs.68647	ESTs, Weakly similar to ALU7_HUMAN ALU SUBFAMILY SQ SEQUENCE CONTAMINATION
10		WARNING			
	132131	RC_AA293495	AF069291	Hs.40539	chromosome 8 open reading frame 1
	115536	RC_AA347193	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), actin binding protein
	132411	RC_AA398474_s	AA059412	Hs.47986	hypothetical protein MGC10940
	115575	RC_AA398512	AA393254	Hs.43619	ESTs
15	115601	RC_AA400277	AA148984	Hs.48849	ESTs, Weakly similar to ALU4_HUMAN ALU SUBFAMILY SB2 SEQUENCE CONTAMINATION
		WARNING			
	103928	RC_AA400896	D14540	Hs.199160	myeloid/lymphoid or mixed-lineage leukemia (t(11)hoxa (Drosophila) homolog)
	125819	RC_AA404494	AA044840	Hs.251871	CTP synthase
20	115683	RC_AA410345	AF255910	Hs.54650	junctional adhesion molecule 2
	115715	RC_AA416733	BE395161	Hs.1390	proteasome (prosome, macropain) subunit, beta type, 2
	132952	RC_AA425154	AI658580	Hs.61426	Homo sapiens mesenchymal stem cell protein DSC96 mRNA, partial cds
	115819	RC_AA426573	AA486620	Hs.41135	endomucin-2
	132525	RC_AA431418	AW292809	Hs.50727	N-acetylglucosaminidase, alpha- (Sanfilippo disease IIIB)
	115895	RC_AA436182	AB033035	Hs.51965	KIAA1209 protein
25	132333	RC_AA437099	AA192669	Hs.45032	ESTs
	115962	RC_AA446585	AI636361	Hs.179520	hypothetical protein MGC10702
	115967	RC_AA446887	AI745379	Hs.42911	ESTs
	115974	RC_AA447224	BE513442	Hs.238944	hypothetical protein FLJ10631
	115985	RC_AA447709	AA447709	Hs.288115	ESTs, Weakly similar to T08599 probable transcription factor CA150 [H.sapiens]
30	29254	RC_AA453624	AA252468	Hs.1098	DKFZp434J1813 protein
	133071	RC_AA455044	BE384932	Hs.64313	ESTs, Weakly similar to AF257182 1 G-protein-coupled receptor 48 [H.sapiens]
	116095	RC_AA456045	AA043428	Hs.62618	ESTs
	122691	RC_AA460454_s	R19768	Hs.172788	ALEX3 protein
	116210	RC_AA476494	BE622792	Hs.172788	ALEX3 protein
35	116213	RC_AA476738	AA292105	Hs.326740	hypothetical protein MGC10947
	134585	RC_AA481422	D14041	Hs.278573	H-2K binding factor-2
	134790	RC_AA482269	BE002798	Hs.287850	integral membrane protein 1
	116265	RC_AA482595	BE297412	Hs.55189	hypothetical protein
	129334	RC_AA485084_s	AW157022	Hs.4947	hypothetical protein FLJ22584
40	116274	RC_AA485431_s	AI129767	Hs.182874	guanine nucleotide binding protein (G protein) alpha 12
	303150	RC_AA489057	AA887146	Hs.8217	stromal antigen 2
	129945	RC_AA489638	BE514376	Hs.165998	PAI-1 mRNA-binding protein
	116331	RC_AA491000	N41300	Hs.71616	Homo sapiens mRNA; cDNA DKFZp586N1720 (from clone DKFZp586N1720)
	116333	RC_AA491250	AF155827	Hs.203963	hypothetical protein FLJ10339
45	132994	RC_AA505133	AA112748	Hs.279905	clone HQ0310 PRO0310p1
	134577	RC_AA509447	BE244323	Hs.85951	exportin, tRNA (nuclear export receptor for tRNAs)
	116391	RC_AA509423	T86558	Hs.75113	general transcription factor IIIA
	116394	RC_AA509574_J	NM_006033	Hs.65370	lipase, endothelial
	134531	RC_AA600153	AI742845	Hs.110713	DEK oncogene (DNA binding)
50	116417	RC_AA609309	AW499664	Hs.12484	Human clone 23826 mRNA sequence
	116429	RC_AA609710	AF191018	Hs.279923	putative nucleotide binding protein, estradiol-induced
	116439	RC_AA610068	AA251594	Hs.43913	PIBF1 gene product
	116459	RC_AA621399	R80137	Hs.302738	Homo sapiens cDNA: FLJ21425 fls, clone COL04162
	427505	RC_AA621752	AA361562	Hs.178761	26S proteasome-associated pad1 homolog
55	132699	RC_C21523	AW449822	Hs.55200	ESTs
	116541	RC_D12160	D12160	Hs.249212	polymerase (RNA) III (DNA directed) (155kD)
	132557	RC_D19708	AA114926	Hs.5122	ESTs
	112259	RC_D25801	AA337548	Hs.333402	hypothetical protein MGC12760
60	116571	RC_D45652	D45652		gb:U00802848 Human adult lung 3' directed Mbol cDNA Homo sapiens cDNA 3', mRNA
		sequence,			
	129815	RC_D60208_f	BE565817	Hs.26498	hypothetical protein FLJ21657
	421919	RC_D80504_s	AJ224901	Hs.109526	zinc finger protein 198
	116643	RC_F03010	AI367044	Hs.153638	myeloid/lymphoid or mixed-lineage leukemia 2
	116661	RC_F04247	R61504		gb:Y16a03.s1 Soares infant brain 1N1B Homo sapiens cDNA clone 3' similar to contains Alu
65		repetitive			
	116715	RC_F10966	AL117440	Hs.170263	tumor protein p53-binding protein, 1
	116729	RC_F13700	BE549407	Hs.115823	ribonuclease P, 40kD subunit
	318709	RC_H05063	R52576	Hs.285280	Homo sapiens cDNA: FLJ22096 fls, clone HEP16953
	134760	RC_H16758	NM_000121	Hs.89548	erythropoietin receptor
70	116773	RC_H17315_s	AI823410	Hs.169149	karyopherin alpha 1 (importin alpha 5)
	106425	RC_H22556	H24201	Hs.247423	adducin 2 (beta)
	116780	RC_H22566	H22566	Hs.30098	ESTs
	131978	RC_H48459_s	AA355925	Hs.36232	KIAA0186 gene product
	116819	RC_H53073	H53073	Hs.93698	EST
75	111428	RC_H55559_s	AL031428	Hs.174174	KIAA0601 protein
	133175	RC_H57957_s	AW955632	Hs.66666	ESTs, Weakly similar to S19560 proline-rich protein MP4 - mouse [M.musculus]

5	116844	RC_H64938_s	H64938	Hs.337434	ESTs, Weakly similar to A46010 X-linked retinopathy protein [H.sapiens]
	116845	RC_H64973	AA649530		gb:ns4405.s1 NCL CGAP_Alv1 Homo sapiens cDNA clone, mRNA sequence
	116892	RC_H69535	AI573283	Hs.38458	ESTs
	116925	RC_H73110	H73110	Hs.260603	ESTs, Moderately similar to A47582 B-cell growth factor precursor [H.sapiens]
	116981	RC_H81783	N29218	Hs.40290	ESTs
10	131768	RC_H86259	AC005757	Hs.31809	hypothetical protein
	117031	RC_H88353	H88353		gb:yyw21a02.s1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:252842 3' similar to contains L1
	117034	RC_H88639	U72209	Hs.180324	YY1-associated factor 2
	132542	RC_H88675	AL137751	Hs.263671	Homo sapiens mRNA; cDNA DKFZp434i0812 (from clone DKFZp434i0812); partial cds
	134403	RC_H93708_s	AA334551	Hs.82767	sperm specific antigen 2
15	117280	RC_N22107	M18217	Hs.172129	Homo sapiens cDNA: FLJ21409 fls, clone COL03924
	117344	RC_N24046	R19085	Hs.210706	Homo sapiens cDNA: FLJ13182 fls, clone NT2RP3004070
	117422	RC_N27028	AI355562	Hs.43880	ESTs, Weakly similar to A46010 X-linked retinopathy protein [H.sapiens]
	117475	RC_N30205	N30205	Hs.93740	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	117487	RC_N30821	N30621	Hs.44203	ESTs
20	130207	RC_N33258	AF044209	Hs.144904	nuclear receptor co-repressor 1
	117549	RC_N33390	N33390	Hs.44483	EST
	117683	RC_N40180	N40180		gb:yy44d02.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone
	IMAGE:276387	3' similar to			
	117710	RC_N45198	N45198	Hs.47248	ESTs, Highly similar to similar to Cdc14B1 phosphatase [H.sapiens]
25	104514	RC_N45979_s	AF164622	Hs.182982	golgin-67
	117791	RC_N48325	N48325	Hs.93956	EST
	117822	RC_N48913	AA706282	Hs.93963	ESTs
	129547	RC_N49394	AB018259	Hs.118140	KIAA0716 gene product
	117895	RC_N50856	AW450348	Hs.93996	ESTs, Highly similar to SORL1_HUMAN SORTILIN-RELATED RECEPTOR PRECURSOR
30	131557	RC_N50721	AA317439	Hs.28707	signal sequence receptor, gamma (translocin-associated protein gamma)
	133057	RC_N53143	AA465131	Hs.64001	Homo sapiens clone 25218 mRNA sequence
	118103	RC_N55326	AA401733	Hs.184134	ESTs
	118111	RC_N55493	N55493		gb:yy50c02.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:246146 3'
	118129	RC_N57493	N57493		gb:yy54c08.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone
35	IMAGE:277358	3', mRNA			
	118278	RC_N62955	N62955	Hs.316433	Homo sapiens cDNA FLJ11375 fls, clone HEMBA1000411, weakly similar to ANKYRIN
	118329	RC_N63520	N63520		gb:yy62d01.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone IMAGE:278137
	118336	RC_N63604	BE327311	Hs.47166	HT021
	132457	RC_N64166	AB017365	Hs.173859	frizzled (Drosophila) homolog 7
40	118363	RC_N64168	AI183838	Hs.48938	hypothetical protein FLJ21802
	118384	RC_N64191	N46114	Hs.29169	hypothetical protein FLJ22623
	118475	RC_N66845	N66845		gb:za46c11.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:295604 3'
	118491	RC_N67135	AV647908	Hs.90424	Homo sapiens cDNA: FLJ23285 fls, clone HEP09071
	118500	RC_N67295	W32889	Hs.154329	ESTs
45	101663	RC_N68399	NM_003528	Hs.2178	H2B histone family, member Q
	118584	RC_N68963	AW136928		gb:U1-H-B1-adp-d-08-0-U1.s1 NCL CGAP_Sub3 Homo sapiens cDNA clone 3', mRNA
	421983	RC_N69331	AI252640	Hs.110364	peptidylprolyl isomerase C (cyclophilin C)
	118661	RC_N70777	AL137554	Hs.49927	protein kinase NYD-SP15
	118684	RC_N71364_s	N71313	Hs.163986	Homo sapiens cDNA: FLJ22765 fls, clone KIAA1180
50	118689	RC_N71545_s	AW390601	Hs.184544	Homo sapiens, clone IMAGE:3355383, mRNA, partial cds
	118690	RC_N71571	N71571	Hs.269142	ESTs
	118766	RC_N74456	N74456	Hs.50499	EST
	118793	RC_N75594	N75594	Hs.285921	ESTs, Moderately similar to T47135 hypothetical protein DKFZp761L0812.1 [H.sapiens]
	118817	RC_N79035	AI688658	Hs.50797	ESTs
55	118844	RC_N80279	AL035364	Hs.50891	hypothetical protein
	118919	RC_N91797	AW452696	Hs.130760	myosin phosphatase, target subunit 2
	129558	RC_N92454	AW580922	Hs.180446	karyopherin (importin) beta 1
	132692	RC_N94581	AW191962	Hs.249239	collagen, type VIII, alpha 2
	118996	RC_N94746	N94746	Hs.274248	hypothetical protein FLJ20758
60	119021	RC_N98238	N98238	Hs.55185	ESTs
	119039	RC_R02384	AI160570	Hs.252097	pregnancy specific beta-1-glycoprotein 6
	119063	RC_R16833	R16833	Hs.53106	ESTs, Moderately similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
	WARNING				
	118523	RC_R41828_s	Y07759	Hs.170157	myosin VA (heavy polypeptide 12, myosin)
65	119111	RC_R43203	T02865	Hs.328321	EST
	133970	RC_R46395	AA214228	Hs.127751	hypothetical protein
	119146	RC_R58863	R58863	Hs.91815	ESTs
	120296	RC_R78248	AW995911	Hs.299883	hypothetical protein FLJ23399
	119239	RC_T11483	T11483		gb:CHR90049 Chromosome 9 exon Homo sapiens cDNA clone 111-1 5' and 3', mRNA
70	sequence				
	119281	RC_T16896	AI692322	Hs.65373	ESTs, Weakly similar to T02345 hypothetical protein KIAA0324 [H.sapiens]
	119298	RC_T23820	NM_001241	Hs.155478	cyclin T2
	126502	RC_T30222	T10077	Hs.13453	hypothetical protein FLJ14753
	135073	RC_W15275_s	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (from clone DKFZp586E1624)

	119558	RC_W38194	W38194		Empirically selected from AFFX single probeset
	132736	RC_W42414_s	AW081883	Hs.288261	Homo sapiens cDNA: FLJ23037 fs, clone LNG02036, highly similar to HSU68019 Homo sapiens mad protein
5	132173	RC_W46577_s	X89426	Hs.41716	endothelial cell-specific molecule 1
	134873	RC_W49632_s	AA884471	Hs.90449	Human clone 23908 mRNA sequence
	119650	RC_W57613	R82342	Hs.79856	ESTs, Weakly similar to S65657 alpha-1C-adrenergic receptor splice form 2 [H.sapiens]
	119654	RC_W57759	W57759		gb:zd20g11.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:341252 3' similar to
10	119683	RC_W61118	W65379	Hs.57835	ESTs
	119694	RC_W65344	AA041350	Hs.57847	ESTs, Moderately similar to ICE4_HUMAN CASPASE-4 PRECURSOR [H.sapiens]
	119718	RC_W69216	W69216	Hs.92848	ESTs
	133010	RC_W69379	AI287518	Hs.62669	Homo sapiens mRNA; cDNA DKFZp586D0923 (from clone DKFZp586D0923)
	119938	RC_W86728	AW014862	Hs.58885	ESTs
15	120128	RC_Z38499	BE379320	Hs.91448	MKP-1 like protein tyrosine phosphatase
	120130	RC_Z38630	AA045767	Hs.5300	bladder cancer associated protein
	120148	RC_Z39494	F02806	Hs.65765	ESTs
	120155	RC_Z39623	Z39623	Hs.65783	ESTs
	131486	RC_Z40071_s	F06972	Hs.27372	BMX non-receptor tyrosine kinase
20	120183	RC_Z40174	AW082866	Hs.65882	ESTs
	120184	RC_Z40182	Z40182	Hs.65885	EST
	120211	RC_Z40904	Z40904	Hs.66012	EST
	120245	RC_AA166965	AW959615	Hs.111045	ESTs
	120247	RC_AA167500	AA167500	Hs.103939	EST
25	120254	RC_AA169599_s	W50403	Hs.111054	ESTs
	120259	RC_AA171724	AW014786	Hs.192742	hypothetical protein FLJ12785
	120260	RC_AA171739	AK000061	Hs.101590	hypothetical protein
	120275	RC_AA177105	AA177105	Hs.78457	solute carrier family 25 (mitochondrial carrier, ornithine transporter) member 15
	120284	RC_AA182626	AA179656		gb:zp54e11.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 3' similar to contains
30	114056	RC_AA186324	AA188175	Hs.82506	KIAA1254 protein
	129507	RC_AA192099	AJ236885	Hs.112180	zinc finger protein 148 (pH2-52)
	120302	RC_AA192173	AA837098	Hs.269933	ESTs
	120303	RC_AA192415	AI216292	Hs.96184	ESTs
35	120305	RC_AA192553	AW295096	Hs.101337	uncoupling protein 3 (mitochondrial, proton carrier)
	120319	RC_AA194851	T57776	Hs.191094	ESTs
	133389	RC_AA196520_s	AA195764	Hs.72639	ESTs
	120326	RC_AA196300	AA196300	Hs.21145	hypothetical protein RG083M05.2
	134272	RC_AA196517	X76040	Hs.278614	protease, serine, 15
40	133145	RC_AA196549	H94227	Hs.6592	Homo sapiens, clone IMAGE:2961368, mRNA, partial cds
	120327	RC_AA196721	AK000292	Hs.278732	hypothetical protein FLJ20285
	106686	RC_AA196729_s	N66397	Hs.334825	Homo sapiens cDNA FLJ14752 fs, clone NT2RP3003071
	120328	RC_AA196979	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapiens]
	120340	RC_AA206828	AA206828		gb:zq80b08.s1 Stratagene hNT neuron (937233) Homo sapiens cDNA clone IMAGE:647895 3' similar to
45	134292	RC_AA207123	AI906291	Hs.81234	immunoglobulin superfamily, member 3
	131522	RC_AA214539_s	AI380040	Hs.239489	TIAT1 cytotoxic granule-associated RNA-binding protein
	129051	RC_AA226914_s	AA227068	Hs.108301	nuclear receptor subfamily 2, group C, member 1
	120375	RC_AA227260	AF028706	Hs.111227	Zic family member 3 (odd-paired Drosophila homolog, heterotaxy 1)
50	120376	RC_AA227469	AA227469		gb:zr18a07.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:663732 3', mRNA sequence.
	120390	RC_AA233122	AA837093	Hs.111460	calcium/calmodulin-dependent protein kinase (CaM kinase) II delta
	303876	RC_AA233334_s	U84820	Hs.66521	Machado-Joseph disease (spinocerebellar ataxia 3, olivopontocerebellar ataxia 3, autosomal dominant, ataxin 3)
55	132038	RC_AA233347	AI825842	Hs.3776	zinc finger protein 216
	104463	RC_AA233519	T85825	Hs.246885	hypothetical protein FLJ20783
	125750	RC_AA233714	AA018515	Hs.264482	Homo sapiens mRNA; cDNA DKFZp761A0411 (from clone DKFZp761A0411)
	120396	RC_AA233796	AA134006	Hs.79306	eukaryotic translation initiation factor 4E
	120409	RC_AA235050_f	AA235050		gb:zs38e04.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:687486 3' similar to gb:1.07077
60	120414	RC_AA235704	AW137156	Hs.181202	hypothetical protein FLJ10038
	120420	RC_AA236031	AI128114	Hs.112885	spinal cord-derived growth factor-B
	120422	RC_AA236352	AL133097	Hs.301717	hypothetical protein DKFZp434N1928
	132221	RC_AA236390_s	W94915	Hs.42419	ESTs
65	120423	RC_AA236453	AA238453	Hs.18978	Homo sapiens cDNA: FLJ22822 fs, clone KIAA3968
	120435	RC_AA243370	AA243370	Hs.96450	EST
	120453	RC_AA250947	AA250947	Hs.170263	tumor protein p53-binding protein, 1
	120455	RC_AA251083	AA251720	Hs.104347	ESTs, Weakly similar to ALLC_HUMAN IIII ALU CLASS C WARNING ENTRY III [H.sapiens]
	120456	RC_AA251113	AA488750	Hs.88414	BTB and CNC homology 1, basic leucine zipper transcription factor 2
70	120473	RC_AA251973	AA251973	Hs.269988	ESTs
	128922	RC_AA252023	AI244901	Hs.9589	ubiquitin 1
	120477	RC_AA252414	AA252414	Hs.43141	DKFZP727C091 protein
	120479	RC_AA252650	AF006689	Hs.110299	mitogen-activated protein kinase kinase 7
	120488	RC_AA255523	AW952916	Hs.63510	KIAA0141 gene product
75	120510	RC_AA258128	AI796395	Hs.111377	ESTs
	120527	RC_AA262105	AA262105	Hs.4094	Homo sapiens cDNA FLJ14208 fs, clone NT2RP3003264
	120528	RC_AA262107	AI923511	Hs.104413	ESTs

	120529	RC_AA262235	AI434823	Hs.104415	ESTs
	120541	RC_AA278298	W07318	Hs.240	M-phase phosphoprotein 1
	131445	RC_AA278529_J	NM_014264	Hs.172052	serine/threonine kinase 18
	120544	RC_AA278721	BE548277	Hs.103104	ESTs
5	120562	RC_AA280036	BE244580	Hs.302267	hypothetical protein FLJ10330
	120569	RC_AA280648	AA807544	Hs.24970	ESTs, Weakly similar to B34323 GTP-binding protein Rab2 [H.sapiens]
	120571	RC_AA280738	AB037744	Hs.34892	KIAA1323 protein
	120572	RC_AA280794	H39599	Hs.294008	ESTs
	129434	RC_AA280837	AW967495	Hs.186644	ESTs
10	130529	RC_AA280886	AA178953		gb:zfp39e03.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 3' similar to contains Alu repetitive
	120575	RC_AA280934	AW978022	Hs.238911	hypothetical protein DKFZp762E1511; KIAA1816 protein
	132635	RC_AA281535	AB020686	Hs.54037	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)
	120591	RC_AA281797_s	AF078847	Hs.191356	general transcription factor IIF, polypeptide 2 (44kD subunit)
15	120593	RC_AA282047	AA748355	Hs.193522	ESTs
	430275	RC_AA283002	Z11773	Hs.237786	zinc finger protein 187
	117729	RC_AA283709	AA306166	Hs.7145	calpain 7
	120609	RC_AA283902	AW978721	Hs.266076	ESTs, Weakly similar to A46010 X-linked retinopathy protein [H.sapiens]
	132754	RC_AA284108	AI752244	Hs.75309	eukaryotic translation elongation factor 2
20	130315	RC_AA284109	AI241084	Hs.154353	nonselective sodium potassium/proton exchanger
	132814	RC_AA284371	AA284371	Hs.118064	similar to rat nuclear ubiquitous casein kinase 2
	447503	RC_AA284744_J	AA115496	Hs.336898	Homo sapiens, Similar to RIKEN cDNA 1810038N03 gene, clone MGC:9890, mRNA, complete cds
	135376	RC_AA284784	BE617856	Hs.99756	mitochondrial ribosome recycling factor
25	120621	RC_AA284840	AW961294	Hs.143818	hypothetical protein FLJ23459
	107868	RC_AA286844	AA286844	Hs.61260	hypothetical protein FLJ13164
	129868	RC_AA287032	AW172431	Hs.13012	ESTs
	120644	RC_AA287038	AI869129	Hs.96616	ESTs
	120660	RC_AA287546	AA286785	Hs.99677	ESTs
30	135370	RC_AA287553_s	BE622187	Hs.99670	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	120661	RC_AA287556	AA287556	Hs.263412	ESTs, Weakly similar to ALU8_HUMAN III ALU CLASS B WARNING ENTRY III [H.sapiens]
	129116	RC_AA287564	AB019494	Hs.225767	IDN3 protein
	131567	RC_AA291015_s	AF015592	Hs.28853	CDC7 (cell division cycle 7, S. cerevisiae, homolog)-like 1
	120699	RC_AA291716	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribosomal protein L13a, cytosolic [H.sapiens]
35	100690	RC_AA291749_s	AA383256	Hs.1657	estrogen receptor 1
	120726	RC_AA293656	AA293655	Hs.97293	ESTs
	120737	RC_AA302430	AL049176	Hs.82223	chordin-like
	120745	RC_AA302809	AA302809		gb:EST10426 Adipose tissue, white I Homo sapiens cDNA 3' end, mRNA sequence.
	135192	RC_AA302820_s	U83993	Hs.321709	purinergic receptor P2X, ligand-gated ion channel, 4
40	120750	RC_AA310499	AI191410	Hs.96693	ESTs, Moderately similar to 2109260A B cell growth factor [H.sapiens]
	120761	RC_AA321890	AA321890	Hs.1265	branched chain keto acid dehydrogenase E1, beta polypeptide (maple syrup urine disease)
	120768	RC_AA340589	AA340589	Hs.104560	EST
	120769	RC_AA340622	AI769467	Hs.96769	ESTs
	135232	RC_AA342457_J	AL038812	Hs.96800	ESTs, Moderately similar to ALU7_HUMAN ALU SUBFAMILY SQ SEQUENCE
45		CONTAMINATION			
	133439	RC_AA342828_s	Z23091	Hs.73734	glycoprotein V (platelet)
	120793	RC_AA342864	AA342864	Hs.96812	ESTs
	120796	RC_AA342973	AI247356	Hs.96820	ESTs
	120809	RC_AA346495	AA346495		gb:EST52657 Fetal heart II Homo sapiens cDNA 3' end similar to EST containing O family
50		repeat, mRNA sequence.			
	132459	RC_AA347573	AL120071	Hs.48998	fibronectin leucine rich transmembrane protein 2
	120825	RC_AA347614	AI280215	Hs.96885	ESTs
	120827	RC_AA347717	AA382525	Hs.132967	Human EST clone 122887 mariner transposon Hsmar1 sequence
	120839	RC_AA348913	AA348913		gb:EST55442 Infant adrenal gland II Homo sapiens cDNA 3' end similar to EST containing Alu
55		repeat, mRNA sequence.			
	120850	RC_AA349647	AA349647	Hs.96927	Homo sapiens cDNA FLJ12573 fls, clone NT2RM4000979
	120852	RC_AA349773	AA349773	Hs.191564	ESTs
	128852	RC_AA350541_s	R40622	Hs.106601	ESTs
	135240	RC_AA357159_J	AA357159	Hs.96986	EST
60	120870	RC_AA357172_J	AA357172	Hs.292581	ESTs, Moderately similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION WARNING
	134637	RC_AA369856_s	U87309	Hs.180941	vacuolar protein sorting 41 (yeast homolog)
	120894	RC_AA370132	AA370132	Hs.97063	ESTs
	131854	RC_AA370472_s	AF229839	Hs.173202	I-kappa-B-interacting Ras-like protein 1
65	120897	RC_AA370867	AA370867	Hs.97079	ESTs, Moderately similar to AF174605 1 F-box protein Fbx25 [H.sapiens]
	120915	RC_AA377296	AL135556	Hs.97104	ESTs
	120935	RC_AA383902	AL048409	Hs.97177	ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION WARNING
	120936	RC_AA385934	AA385934	Hs.97184	EST, Highly similar to (define not available 7499603) [C.elegans]
	120937	RC_AA386255	AA386255	Hs.97186	EST
	120938	RC_AA386260	AA386260	Hs.104632	EST
	129722	RC_AA386266	R20855	Hs.5422	glycoprotein M6B
	120960	RC_AA398014	AA398014	Hs.104684	EST
	120985	RC_AA398222	AI219896	Hs.97592	ESTs
75	120988	RC_AA398235	AA398235	Hs.97631	ESTs

	121008	RC_AA398348	AA398348	Hs.301720	Human DNA sequence from clone RP11-251J8 on chromosome 13 Contains ESTs, STSs, GSSs and a CpG
	121029	RC_AA398482	AA398482	Hs.97641	EST
	121032	RC_AA398504	AA398504	Hs.161788	ESTs
5	121033	RC_AA398505	AA398505	Hs.97360	ESTs
	121034	RC_AA398507	AA398507	Hs.271623	nucleoporin 50kD
	121035	RC_AA398523	AA398523	Hs.210579	ESTs
	121058	RC_AA398625	AA398625	Hs.97391	ESTs
	121060	RC_AA398632	AA398632	Hs.97395	ESTs
10	121061	RC_AA398633	AA398633	Hs.97396	ESTs
	121091	RC_AA398694	AA398694	Hs.97657	ESTs, Moderately similar to ALU8_HUMAN ALU SUBFAMILY SX SEQUENCE
	CONTAMINATION				
	121092	RC_AA398895	AA398895	Hs.97658	EST
	121094	RC_AA398900	AA402505		gb:z62h10.1 Soares_testis_NHT Homo sapiens cDNA clone 5', mRNA sequence
15	121096	RC_AA398904	AA398904	Hs.332690	ESTs
	121115	RC_AA399122	AA398187	Hs.104682	ESTs, Weakly similar to mitochondrial citrate transport protein [H.sapiens]
	121121	RC_AA399371	AA399371	Hs.189095	similar to SALL1 (sal (Drosophila))-like
	121122	RC_AA399373	AI126713	Hs.192233	ESTs, Highly similar to T00337 hypothetical protein KIAA0568 [H.sapiens]
	121125	RC_AA399441	AL042981	Hs.251278	KIAA1201 protein
20	121151	RC_AA399636	AA399636	Hs.143629	ESTs
	121153	RC_AA399640	AA399640	Hs.97694	ESTs
	121163	RC_AA399680	AI676062	Hs.111902	ESTs
	121176	RC_AA400080	AL121523	Hs.97774	ESTs
	121192	RC_AA400262	AA400262	Hs.190093	ESTs
25	121223	RC_AA400725	AI002110	Hs.97169	ESTs, Weakly similar to dJ667H12.2.1 [H.sapiens]
	121227	RC_AA400748	AA400748	Hs.97823	Homo sapiens mRNA; cDNA DKFZp434D024 (from clone DKFZp434D024)
	121231	RC_AA400780	AA814948	Hs.96343	ESTs, Weakly similar to ALUC_HUMAN III ALU CLASS C WARNING ENTRY III [H.sapiens]
	121278	RC_AA401631	AA037121	Hs.98518	Homo sapiens cDNA FLJ11490 fis, clone HEMBA1001918
	121279	RC_AA401688	AA292873	Hs.177996	ESTs
30	121282	RC_AA401695	AA401695	Hs.97334	ESTs
	121299	RC_AA402227	AA402227	Hs.22826	tropomodulin 3 (ubiquitous)
	121301	RC_AA402329	NM_006202	Hs.89901	phosphodiesterase 4A, cAMP-specific (dunce (Drosophila))-homolog phosphodiesterase E2)
	121302	RC_AA402398	AA402587	Hs.325520	LAT1-3TM protein
	121304	RC_AA402449	AA293863	Hs.97316	EST
35	121305	RC_AA402468	AA402468	Hs.291557	ESTs
	134721	RC_AA403268_s	AK000112	Hs.89306	hypothetical protein FLJ20105
	121323	RC_AA403314	AA291411	Hs.97247	ESTs
	121324	RC_AA404229	AA404229	Hs.97842	EST
40	129047	RC_AA404260	AI768623	Hs.108264	ESTs
	131074	RC_AA404271	U16125	Hs.181581	glutamate receptor, ionotropic, kainate 1
	121344	RC_AA405026	AA405026	Hs.193754	ESTs
	121348	RC_AA405182	AA405182	Hs.97973	ESTs
	121350	RC_AA405237	AA405237		gb:z06e10.s1 NC1_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712362 3' similar to contains Alu
45	121400	RC_AA406061	AA406061	Hs.98001	EST
	121402	RC_AA406063	AA406063	Hs.98003	ESTs
	121403	RC_AA406070	AA406070	Hs.98004	EST
	121408	RC_AA406137	AA406137	Hs.98019	EST
	121431	RC_AA406335	AA035279	Hs.176731	ESTs
50	132936	RC_AA411801	AL120659	Hs.6111	aryl-hydrocarbon receptor nuclear translocator 2
	121471	RC_AA411804	AA411804	Hs.261575	ESTs
	121474	RC_AA411833	AA402335	Hs.188760	ESTs, Highly similar to Trad [H.sapiens]
	121526	RC_AA412219	AW665325	Hs.98120	ESTs
	121530	RC_AA412259	AA778658	Hs.98122	ESTs
55	121558	RC_AA412497	AA412497		gb:z05g12.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730150 3' similar to contains L1.L1
	121559	RC_AA412498	AI192044	Hs.104778	ESTs
	121584	RC_AA416586	AI024471	Hs.98232	ESTs
	121609	RC_AA416867	AA416867	Hs.98185	EST
60	121612	RC_AA416874	AA416874	Hs.98168	ESTs
	121737	RC_AA421133	AA421133	Hs.104671	erythrocyte transmembrane protein
	121740	RC_AA421138	AA421138	Hs.98334	EST
	129194	RC_AA422079	AA150797	Hs.108276	latexin protein
	121784	RC_AA423837	T90789	Hs.94308	RAB35, member RAS oncogene family
65	121802	RC_AA424328	AI251870	Hs.188898	ESTs
	121803	RC_AA424339	AI338371	Hs.157173	ESTs
	135286	RC_AA424469_s	AW023482	Hs.97849	ESTs
	121806	RC_AA424502	AA424313	Hs.98402	ESTs
	129517	RC_AA425004	AW972853	Hs.112237	ESTs
70	121845	RC_AA425734	AI732692	Hs.165006	ESTs, Moderately similar to ALU2_HUMAN ALU SUBFAMILY SB SEQUENCE
	CONTAMINATION				
	121853	RC_AA425887	AA425887	Hs.98502	hypothetical protein FLJ14303
	121891	RC_AA426456	AA426456	Hs.98469	ESTs
	121895	RC_AA427396	AA427396		gb:z033a02.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:771050 3' similar to contains
75	121899	RC_AA427555	R55341	Hs.50421	KIAA0203 gene product

	121917	RC_AA428218	AA406397	Hs.98038	ESTs
	121918	RC_AA428242	BE274689	Hs.184175	chromosome 2 open reading frame 3
	121919	RC_AA428281	AA428281	Hs.98560	EST
	121941	RC_AA428865	AA428865	Hs.98563	ESTs
5	121942	RC_AA428994	AW452701	Hs.293237	ESTs
	121970	RC_AA429666	AA429666	Hs.98617	EST
	121993	RC_AA430181	AW297880	Hs.98661	ESTs
	134680	RC_AA430184_s	U73524	Hs.87465	ATP/GTP-binding protein
	128753	RC_AA431288_s	AA306478	Hs.95327	CD3D antigen, delta polypeptide (TIT3 complex)
10	122022	RC_AA431293	AA431293	Hs.98716	ESTs, Moderately similar to T42650 hypothetical protein DKFZp434D0215.1 [H.sapiens]
	122050	RC_AA431476	AA453076	Hs.166109	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2
	122051	RC_AA431492	AA431492	Hs.98742	EST
	122055	RC_AA431732	AA431732	Hs.98747	EST
	122105	RC_AA432278	AW241685	Hs.98699	ESTs
15	122125	RC_AA434411	AK000492	Hs.98806	hypothetical protein
	135235	RC_AA435512_s	AW298244	Hs.293507	ESTs
	122162	RC_AA435698	AA628233	Hs.79946	cytochrome P450, subfamily XIX (aromatization of androgens)
	129406	RC_AA435711	AB018255	Hs.111138	KIAA0712 gene product
	318801	RC_AA435815_s	U40763	Hs.77965	peptidyl-prolyl isomerase G (cyclophilin G)
20	122186	RC_AA435842	AA398811	Hs.104673	ESTs
	122235	RC_AA436475	AA436475	Hs.112227	membrane-associated nucleic acid binding protein
	129131	RC_AA436489	AB026436	Hs.177534	dual specificity phosphatase 10
	134664	RC_AA442060	AA256106	Hs.87507	ESTs
	122310	RC_AA442079	AW192803	Hs.98974	ESTs, Weakly similar to S65824 reverse transcriptase homolog [H.sapiens]
25	122334	RC_AA443151	BE465894	Hs.98365	ESTs, Weakly similar to LB4D_HUMAN NADP-DEPENDENT LEUKOTRIENE B4 12-
	122382	RC_AA446133	AA446440	Hs.98643	ESTs
	122425	RC_AA447145	AB007859	Hs.100955	KIAA0399 protein
	122431	RC_AA447398	AA447398	Hs.99104	ESTs
	122450	RC_AA447643	AA447643	Hs.112095	hypothetical protein DKFZp434F1619
30	302653	RC_AA447742_s	AJ404468	Hs.284259	dynein, axonemal, heavy polypeptide 9
	122477	RC_AA448226	AA448226	Hs.324123	ESTs
	122500	RC_AA448825	AA448825	Hs.99190	ESTs
	122522	RC_AA449444	AA299607	Hs.98969	ESTs
	122536	RC_AA450087	AF060877	Hs.99236	regulator of G-protein signalling 20
35	122538	RC_AA450211	AA450211	Hs.99239	ESTs
	122540	RC_AA450244	AA476741	Hs.98279	ESTs, Weakly similar to A43932 mucin 2 precursor, intestinal [H.sapiens]
	122560	RC_AA452123	AW392342	Hs.283077	centrosomal P4.1-associated protein; uncharacterized bone marrow protein BM032
	421919	RC_AA452155	AJ224901	Hs.109526	zinc finger protein 198
	122562	RC_AA452156	AA452156		gb:zx29c03.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:787876 3',
40		mRNA			
	122585	RC_AA453036	AI681654	Hs.170737	hypothetical protein FLJ23251
	122608	RC_AA453526	AA453525	Hs.143077	ESTs
	122635	RC_AA454085	AA454085		gb:zx33a08.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:788246 3'
45		similar to			
	122636	RC_AA454103	AW651706	Hs.99519	hypothetical protein FLJ14007
	122653	RC_AA454642	AW009166	Hs.99376	ESTs
	122660	RC_AA454935	AI816827	Hs.180069	nuclear respiratory factor 1
	122703	RC_AA456323	AA456323	Hs.289369	ESTs
	122724	RC_AA457395	AA457395	Hs.99457	ESTs
50	122749	RC_AA458850	AA458850	Hs.293372	ESTs, Weakly similar to B34087 hypothetical protein [H.sapiens]
	122772	RC_AA459662	AW117452	Hs.99489	ESTs
	131098	RC_AA459668	U66669	Hs.236842	3-hydroxyisobutyryl-Coenzyme A hydrolase
	129045	RC_AA459679_s	AI082883	Hs.30732	hypothetical protein FLJ13409; KIAA1711 protein
	122777	RC_AA459702	AK001022	Hs.214397	hypothetical protein FLJ10160 similar to insulin related protein 2
55	135362	RC_AA460017_s	AA978128	Hs.99513	ESTs, Weakly similar to T17454 diaphanous-related formin - mouse [M.musculus]
	122798	RC_AA460324	AW366286	Hs.145696	splicing factor (CC1.3)
	122837	RC_AA461509	AA461509	Hs.293565	ESTs, Weakly similar to putative p150 [H.sapiens]
	122860	RC_AA464414_s	AA464414		gb:zx78g01.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:809904 3',
60		mRNA sequence			
	122861	RC_AA464428	AA335721	Hs.119394	ESTs
	122910	RC_AA470084	AA470084	Hs.98358	ESTs
	132899	RC_AA476606_s	AA476606	Hs.59666	SMAD in the antisense orientation
	122967	RC_AA478521	AA806187	Hs.289101	glucose regulated protein, 58kD
	129560	RC_AA478523	AA317841	Hs.7845	hypothetical protein MGC2752
65	123009	RC_AA479949	AA535244	Hs.78305	RAB2, member RAS oncogene family
	128917	RC_AA481252	AI365215	Hs.206087	oncogene TC21
	123081	RC_AA485351	AI815486	Hs.243901	Homo sapiens cDNA FLJ20738 fs, clone HEP08257
	123133	RC_AA487264	AA487264	Hs.154974	Homo sapiens mRNA; cDNA DKFZp667N064 (from clone DKFZp667N064)
	123184	RC_AA489072	BE247767	Hs.18166	KIAA0870 protein
70	128671	RC_AA489630	NM_014700	Hs.119004	KIAA0665 gene product
	123233	RC_AA490225	AW974175	Hs.188751	ESTs, Weakly similar to MAPB_HUMAN MICROTUBULE-ASSOCIATED PROTEIN 18 [H.sapiens]
	123234	RC_AA490227	NM_001938	Hs.16697	down-regulator of transcription 1, TBP-binding (negative cofactor 2)
	123236	RC_AA490255	AW968504	Hs.123073	CDC2-related protein kinase 7
75	123255	RC_AA490890	AA830335	Hs.105273	ESTs
	129503	RC_AA490916_s	AW768399	Hs.112157	ESTs

	131043	RC_AA490325	AF084535	Hs.22464	epilepsy, progressive myoclonus type 2, Lafora disease (taforin)
	123259	RC_AA490955	AF744152	Hs.283374	ESTs, Weakly similar to CA15_HUMAN COLLAGEN ALPHA 1(V) CHAIN PRECURSOR
		[H.sapiens]			
5	123284	RC_AA495812	AA488988	Hs.293796	ESTs
	123286	RC_AA495824	AA495824	Hs.188822	ESTs, Weakly similar to A46010 X-linked retinopathy protein [H.sapiens]
	123315	RC_AA496369	AA496369		gbzv37d10.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:755827 3' similar to contains
	129179	RC_AA504125_s	AW969025	Hs.109154	ESTs
10	131612	RC_AA521473	AU076668	Hs.334884	SEC10 (S. cerevisiae)-like 1
	123421	RC_AA598440	AA598440	Hs.291154	EST, Weakly similar to I38022 hypothetical protein [H.sapiens]
	123449	RC_AA598899_j	AL049325	Hs.112493	Homo sapiens mRNA; cDNA DKFZp564D036 (from clone DKFZp564D036)
	123021	RC_AA599244	AL044675	Hs.173081	KIAA0530 protein
	132830	RC_AA599694_s	NM_014777	Hs.57730	KIAA0133 gene product
	123497	RC_AA600037	AA765256	Hs.135191	ESTs, Weakly similar to unnamed protein product [H.sapiens]
15	123604	RC_AA609135	AA609135	Hs.293076	ESTs
	129539	RC_AA609582	T47614	Hs.323022	ESTs, Highly similar to p60 katanin [H.sapiens]
	123712	RC_AA609684	AA609684	Hs.112748	Homo sapiens cDNA: FLJ21543 fis, clone COL06171
	123731	RC_AA609839	AA609839		gbz662D1.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone IMAGE:951481 3' similar to
20	130725	RC_AA609862	T98807	Hs.80248	RNA-binding protein gene with multiple splicing
	123800	RC_AA620423	AA620423	Hs.112862	EST
	123841	RC_AA620747	AA620747	Hs.112896	ESTs
	123929	RC_AA621364	AA621364	Hs.112981	ESTs
25	123978	RC_G20653	T98832	Hs.170278	ESTs
	133184	RC_D20085	AA001021	Hs.6685	thyroid hormone receptor interactor 8
	132835	RC_D20749	Z83844	Hs.5790	hypothetical protein dJ37E16.5
	132406	RC_D51285_s	AL133731	Hs.4774	Homo sapiens mRNA; cDNA DKFZp761C1712 (from clone DKFZp761C1712)
	128695	RC_D59972_j	NM_003478	Hs.101299	cutlin 5
30	124028	RC_F04112_j	F04112		gbzHSC2JH062 normalized infant brain cDNA Homo sapiens cDNA clone c-2jh06 3', mRNA sequence.
	124057	RC_F13604	AA902384	Hs.73853	bone morphogenetic protein 2
	134899	RC_H01662	AI609045	Hs.321775	hypothetical protein DKFZp434D1428
	130973	RC_H05135_j	AI638418	Hs.78580	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1
35	124106	RC_H12245	H12245		gbym17a12.r1 Soares infant brain 1N1B Homo sapiens cDNA clone 3', mRNA sequence
	124136	RC_H22842	H22842	Hs.101770	EST
	124165	RC_H30894	H30039	Hs.107674	ESTs
	131229	RC_H43442_s	NM_015340	Hs.2450	leucyl-tRNA synthetase, mitochondrial
	124178	RC_H45996	BE463721	Hs.97101	putative G protein-coupled receptor
40	128948	RC_H69281_j	AI537162	Hs.263988	ESTs
	134374	RC_H69485_j	N22687	Hs.8236	ESTs
	124254	RC_H69899	H69899		gbzyu70c12.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone IMAGE:239158 3' similar to
	129056	RC_H70627_s	AI769958	Hs.108336	ESTs, Weakly similar to ALUE_HUMAN III ALU CLASS E WARNING ENTRY III [H.sapiens]
45	100919	RC_H73050_s	X54534	Hs.278994	Rhesus blood group, CcEa antigens
	130724	RC_H73260	AK001507	Hs.306084	Homo sapiens clone FLB6914 PRO1821 mRNA, complete cds
	100716	RC_H77531_s	X89887	Hs.172350	HIR (histone cell cycle regulation defective, S. cerevisiae) homolog A
	124274	RC_H80552	H80552	Hs.102243	EST
	129078	RC_H80737_s	AI351010	Hs.102267	lysosomal
50	124828	RC_H93412	AW952124	Hs.13094	presenilins associated rhomboid-like protein
	124315	RC_H94892_s	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene homolog A (ras related)
	100747	RC_H95643_s	X04588	Hs.85844	neurotrophic tyrosine kinase, receptor, type 1
	124324	RC_H96552	H96552	Hs.159472	Homo sapiens cDNA: FLJ22224 fis, clone HRC01703
	452933	RC_H97146	AW391423	Hs.288555	Homo sapiens cDNA: FLJ22425 fis, clone HRC08686
	132231	RC_H99131_s	AA662910	Hs.42635	hypothetical protein DKFZp434K2435
55	129170	RC_H99462_s	AW250380	Hs.109059	mitochondrial ribosomal protein L12
	133143	RC_H99837_s	AA094538	Hs.272808	putative transcription regulation nuclear protein; KIAA1689 protein
	132963	RC_N22140	AA099693	Hs.34851	epsilon-tubulin
	135297	RC_N22197	AL118782	Hs.300208	Sec23-interacting protein p125
60	134347	RC_N23756_s	AF164142	Hs.82042	solute carrier family 23 (nucleobase transporters), member 1
	130385	RC_N24134	W56119	Hs.155103	eukaryotic translation initiation factor 1A, Y chromosome
	421642	RC_N24195	AF172066	Hs.106346	retinoic acid repressible protein
	439311	RC_N26739	BE270668	Hs.151945	mitochondrial ribosomal protein L43
	124383	RC_N27098	N27098	Hs.102463	EST
	124387	RC_N27637	N27637	Hs.109019	ESTs
65	129341	RC_N33090	AI193519	Hs.226396	hypothetical protein FLJ11126
	129081	RC_N35967	AI364933	Hs.168913	serine/threonine kinase 24 (Sle20, yeast homolog)
	102827	RC_N38959_j	BE244588	Hs.6456	chaperonin containing TCP1, subunit 2 (beta)
	124433	RC_N39069	AA280319	Hs.288840	PRO1575 protein
	124441	RC_N46441	AW450481	Hs.161333	ESTs
70	132338	RC_N48270_j	AA353868	Hs.182982	golgin-67
	131403	RC_N48365_s	AF473114	Hs.26455	ESTs
	124466	RC_N51316	R10084	Hs.113319	kinesin heavy chain member 2
	132210	RC_N51499_s	NM_007203	Hs.42322	A kinase (PRKA) anchor protein 2
	124483	RC_N53976	AI821780	Hs.179864	ESTs
75	124484	RC_N54157	H66118	Hs.285520	ESTs, Weakly similar to 2109260A B cell growth factor [H.sapiens]
	124485	RC_N54300	AB040833	Hs.15420	KIAA1500 protein

	124494	RC_N54831	N54831	Hs.271381	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	129200	RC_N59849	N59849	Hs.13565	Sam68-like phosphotyrosine protein, T-STAR
	124527	RC_N62132	N79264	Hs.269104	ESTs
	124532	RC_N62375	N62375	Hs.102731	EST
5	133213	RC_N63138	AA903424	Hs.6786	ESTs
	124539	RC_N63172	D54120	Hs.148409	cell division cycle 42 (GTP-binding protein, 25kD)
	133651	RC_N63772	A1301740	Hs.173381	dihydropyrimidinase-like 2
	129196	RC_N63787	BE296313	Hs.265592	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	124575	RC_N68168	N68168		gbza11c01.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 3', mRNA sequence
10	124576	RC_N68201	N68201	Hs.269124	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	124577	RC_N68300	N68300		gbza12g07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:292380 3', mRNA
	124578	RC_N68321	N68321	Hs.231500	EST
	124583	RC_N69575	N69575	Hs.102788	ESTs
15	128501	RC_N75007	AL133572	Hs.199009	protein containing CXXC domain 2
	105691	RC_N75542	AI680737	Hs.289068	Homo sapiens cDNA FLJ11918 f1s, clone HEMBB1000272
	128473	RC_N90066	T78277	Hs.100293	O-linked N-acetylglucosamine (GlcNAc) transferase (UDP-N-acetylglucosamine:polypeptide-N-CGI-47 protein
	128639	RC_N91246	AW582962	Hs.102897	regulator of nonsense transcripts 2; DKFZP434D222 protein
	124652	RC_N92751	W19407	Hs.3862	KIAA0318 protein
20	133137	RC_N93214_s	AB002316	Hs.65746	Homo sapiens cDNA FLJ10495 f1s, clone NT2RP2000287, moderately similar to ZINC FINGER
	124671	RC_N99148	AK001357	Hs.102951	ESTs, Weakly similar to T27173 hypothetical protein Y54G11A.9 - Caenorhabditis elegans
	133054	RC_R07876	AA464836	Hs.291079	ESTs, Weakly similar to T27173 hypothetical protein Y54G11A.9 - Caenorhabditis elegans
	130410	RC_R10865_f	J00077	Hs.155421	alpha-fetoprotein
25	124720	RC_R11056	R05283		gb:ye91c08.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:125102 3' similar to
	124722	RC_R11488	T97733	Hs.185685	ESTs
	129561	RC_R22947	R23053		gbyh31a05.r1 Soares placenta Nb2HP Homo sapiens cDNA clone 5' similar to contains L1
30	132965	RC_R26589_f	RC_R23930_s	AL137586	Hs.52763 anaphase-promoting complex subunit 7
	133740	RC_R37588_s	AI248173	Hs.191460	hypothetical protein MGC12936
	133074	RC_R37613	AW162919	Hs.170160	RAB2, member RAS oncogene family-like
	124757	RC_R38398	AL134275	Hs.6434	hypothetical protein DKFZp761F2014
35	124762	RC_R39179_f	H11368	Hs.141055	Homo sapiens clone 23758 mRNA sequence
	124773	RC_R40923	AA553722	Hs.92096	ESTs, Moderately similar to A46010 X-linked retinopathy protein [H.sapiens]
	135266	RC_R41179	R45154	Hs.106604	ESTs
	131375	RC_R41294_s	R41179	Hs.97393	KIAA0328 protein
	133753	RC_R42307_f	AW293165	Hs.143134	ESTs
40	128540	RC_R43189_f	NM_004427	Hs.165263	early development regulator 2 (homolog of polyhomeotic 2)
	124785	RC_R43306	AW297929	Hs.328317	EST
	124792	RC_R44357	W38537	Hs.280740	hypothetical protein MGC3040
	124793	RC_R44519	R44357	Hs.48712	hypothetical protein FLJ20736
45	124799	RC_R45088	R44519		gb:yg24h04.s1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:33350 3', mRNA sequence.
	124812	RC_R47948_j	R45088		gb:yg38g04.s1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:34898 3', mRNA sequence.
	124821	RC_R51524	R47948	Hs.188732	ESTs
	127274	RC_R54950	H87832	Hs.7388	kelch (Drosophila)-like 3
50	124835	RC_R55241	AW966158	Hs.58582	Homo sapiens cDNA FLJ12789 f1s, clone NT2RP2001947
	124845	RC_R59585	R55241	Hs.101214	EST
	124847	RC_R60044	R59585	Hs.101255	ESTs
	440630	RC_R60872	W07701	Hs.304177	Homo sapiens clone FLB8503 PRO2286 mRNA, complete cds
55	124861	RC_R66690	BE561430	Hs.239388	Human DNA sequence from clone RP1-304B14 on chromosome 6. Contains a gene for a novel protein and a part of a gene for a novel protein with two isoforms. Contains ESTs, STSs, GSSs and a CpG island
	130141	RC_R67266_s	R67567	Hs.107110	ESTs
	124879	RC_R73588	NM_004455	Hs.150956	exostoses (multiple)-like 1
	124892	RC_R79403	R73588	Hs.101533	ESTs
60	124906	RC_R87647	AI970003	Hs.23756	hypothetical protein similar to swine acylneuraminase lyase
	124922	RC_R93622	H75964	Hs.107815	ESTs
	124940	RC_R99599_s	R93622	Hs.12163	eukaryotic translation initiation factor 2, subunit 2 (beta, 38kD)
	124941	RC_R99612	AF068846	Hs.103804	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)
	124943	RC_T02888	AI766651	Hs.27774	ESTs, Highly similar to AF161349 1 HSPC086 [H.sapiens]
65	124947	RC_T03170	AW963279	Hs.123373	ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
	124954	RC_T10465	T03170	Hs.100165	ESTs
	132924	RC_T15418_f	AW964237	Hs.6728	KIAA1548 protein
	133113	RC_T15597_f	U55184	Hs.154145	hypothetical protein FLJ11585
70	132975	RC_T15652_j	BE383768	Hs.65238	95 kDa retinoblastoma protein binding protein; KIAA0661 gene product
	131082	RC_T26544_j	R43504	Hs.6181	ESTs
	124980	RC_T40841	AW960782	Hs.6856	ash2 (absent, small, or homeotic, Drosophila, homolog)-like
	124984	RC_T47566_j	AI091121	Hs.246218	Homo sapiens cDNA: FLJ21781 f1s, clone HEP00223
	124991	RC_T50116	T40841	Hs.98681	ESTs
75	129475	RC_T50145_s	BE313210	Hs.223241	eukaryotic translation elongation factor 1 delta (guanine nucleotide exchange protein)
			T50116		gbyb77c10.s1 Stratagene ovary (937217) Homo sapiens cDNA clone IMAGE:77202 3' similar to similar to SP:VE22_LAMBD P03756 EA22 GENE, mRNA sequence.
			NM_004477	Hs.203772	FSHD region gene 1

	125000	RC_T58615	T58615	Hs.110640	ESTs
	132932	RC_T59940_f	AW118826	Hs.6093	Homo sapiens cDNA: FLJ22783 fis, clone KAlA1993
	129534	RC_T63595	AK002126	Hs.11260	hypothetical protein FLJ11264
	125008	RC_T64891	T91251		gb y60a10.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 3', mRNA sequence
5	125009	RC_T64924	T64924	Hs.303046	ESTs
	132940	RC_T64933_r	T79136	Hs.127243	Homo sapiens mRNA for KIAA1724 protein, partial cds
	125017	RC_T68875	T68875		gb yc3005.s1 Stratagene liver (937224) Homo sapiens cDNA clone IMAGE:82209 3', mRNA sequence.
	125018	RC_T69027	T69027	Hs.57475	sex comb on midleg homolog 1
10	125020	RC_T69924	T69981		gb yc19d03.r1 Stratagene lung (937210) Homo sapiens cDNA clone 5', mRNA sequence
	125891	RC_T70353	AI084813	Hs.13197	ESTs
	134204	RC_T79780_s	AI873257	Hs.7994	hypothetical protein FLJ20551
	125050	RC_T79951	AW970209	Hs.111805	ESTs
	125052	RC_T80174_s	T85104	Hs.222779	ESTs, Moderately similar to similar to NEDD-4 [H.sapiens]
15	125054	RC_T80622	T80622	Hs.268601	ESTs, Weakly similar to envelope [H.sapiens]
	125063	RC_T83352	T83352		gb y62d01.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:114721 3' similar to contains Alu repetitive element; contains L1 repetitive element, mRNA sequence.
	125064	RC_T85373	T85373		gb y62d07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:114757 3' similar to contains Alu repetitive element; contains MER3 repetitive element, mRNA sequence.
20	125066	RC_T86284	T86284		gb y67b07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 3' similar to contains Alu repetitive element, mRNA sequence
	112264	RC_T89579_s	AL045364	Hs.79353	transcription factor Dp-1
	125080	RC_T90360	T90360	Hs.268620	ESTs, Highly similar to ALU6_HUMAN ALU SUBFAMILY SP SEQUENCE CONTAMINATION
	125097	RC_T94328_J	AW576389	Hs.335774	EST, Moderately similar to S65657 alpha-1C-adrenergic receptor splice form 2 [H.sapiens]
25	125104	RC_T95590	T95590		gb ye40a03.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 3' similar to gb M10817 GURRAA Iguana Iguana 5S (rRNA), mRNA sequence
	135107	RC_T97257_J	T97257	Hs.337531	ESTs, Moderately similar to I38022 hypothetical protein [H.sapiens]
	129550	RC_T97599_J	AA845462	Hs.124024	deltax (Drosophila) homolog 1
30	125118	RC_T97620	R10606		gb y35f11.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:128877 3' similar to contains Alu repetitive element, mRNA sequence.
	125120	RC_T97775	T97775	Hs.100717	EST
	134160	RC_T98152	T98152	Hs.79432	fibrillin 2 (congenital contractural arachnodactyly)
35	125136	RC_W31479	AW962364	Hs.129051	ESTs
	125144	RC_W37999	AB037742	Hs.24336	KIAA1321 protein
	125150	RC_W38240	W38240		Empirically selected from AFFX single probeset
	104180	RC_W40150	AA247778	Hs.119155	Homo sapiens mRNA full length insert cDNA clone EUROMIMAGE 814975
	131987	RC_W45435	AW453069	Hs.3657	activity-dependent neuroprotective protein
40	125178	RC_W58202	W58127	Hs.31845	ESTs
	125180	RC_W58344	W58469	Hs.103120	ESTs
	125182	RC_W58650	AA451755	Hs.263560	ESTs
	130588	RC_W68736	AL030996	Hs.16411	hypothetical protein LOC57187
	125197	RC_W69106	AF086270	Hs.278554	heterochromatin-like protein 1
	133497	RC_W69111	BE617303	Hs.74266	hypothetical protein MGC4251
45	100562	RC_W69385_s	NM_006185	Hs.301512	nuclear mitotic apparatus protein 1
	125639	RC_W69399_s	Z97630	Hs.226117	H1 histone family, member 0
	129232	RC_W69459	R98881	Hs.109655	sex comb on midleg (Drosophila)-like 1
	101495	RC_W72424	W72424	Hs.112405	S100 calcium-binding protein A9 (calgranulin B)
50	125209	RC_W72724	W72724	Hs.103174	ESTs, Weakly similar to TSP2_HUMAN THROMBOSPONDIN 2 PRECURSOR [H.sapiens]
	125212	RC_W72834	AA746225	Hs.103173	ESTs
	129132	RC_W73955	BE383436	Hs.106847	hypothetical protein MGC2749
	125223	RC_W74701	AI916269	Hs.109057	ESTs, Weakly similar to ALU5_HUMAN ALU SUBFAMILY SC SEQUENCE CONTAMINATION
	125225	RC_W76540	W74169	Hs.16492	DKFZP584G0222 protein
55	125228	RC_W79397	AA033982	Hs.110059	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	132393	RC_W85888	AL135094	Hs.47334	hypothetical protein FLJ14495
	125238	RC_W86038	N99713	Hs.108514	ESTs
	125247	RC_W86881	AA694191	Hs.163914	ESTs
	129296	RC_W87804	AI051967	Hs.110122	ESTs
60	125263	RC_W88942	AA098878		gb zn45g10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 5', mRNA sequence
	125266	RC_W90022	W90022	Hs.186809	ESTs, Highly similar to LCT2_HUMAN LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2
	131321	RC_W92272	U91543	Hs.25601	chromodomain helicase DNA binding protein 3
	131601	RC_W92764_s	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced protein 6
65	131677	RC_W93040	H05317	Hs.283549	ESTs
	120837	RC_W93092	BE149656	Hs.306621	Homo sapiens cDNA FLJ11963 fis, clone HEMBB1001051
	125277	RC_W93227	W93227	Hs.103245	EST
	125278	RC_W93523	AI218439	Hs.129998	enhancer of polycomb 1
70	125280	RC_W93659	AI123705	Hs.106932	ESTs
	131856	RC_W94003_s	W93949	Hs.33245	ESTs
	131844	RC_W94401_s	AI419294	Hs.324342	ESTs
	125284	RC_W94688	NM_002666	Hs.103253	perlepin
	313447	RC_W94787_s	AW016321	Hs.82306	deshrn (actin depolymerizing factor)
	130799	RC_Z38294_s	AB028945	Hs.12696	cortactin SH3 domain-binding protein
75	125289	RC_Z38311	T34530	Hs.4210	Homo sapiens cDNA FLJ13069 fis, clone NT2RP3001752
	128874	RC_Z38465_s	H08245	Hs.106801	ESTs, Weakly similar to PC4259 fertilin associated protein [H.sapiens]

5	130966	RC_Z38525_s	AW971018	Hs.21659	ESTs
	128875	RC_Z38538_f	AB040923	Hs.106808	kelch (Drosophila)-like 1
	133200	RC_Z38551_s	AB037715	Hs.183639	hypothetical protein FLJ10210
	130158	RC_Z38783_s	AB032947	Hs.151301	Ca2+-dependent activator protein for secretion
	125295	RC_Z39113	AB022317	Hs.25887	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4F
10	125298	RC_Z39255_f	AW972542	Hs.289008	Homo sapiens cDNA: FLJ21814 f1s, clone HEP01068
	125300	RC_Z39591	Z39591	Hs.101376	EST
	323122	RC_Z39783_s	BE622770	Hs.264915	Homo sapiens cDNA FLJ12908 f1s, clone NT2RP2004399
	311463	RC_Z39920	R55344	Hs.22142	cytochrome b5 reductase b5R.2
	130882	RC_Z40166_f	AA497044	Hs.20887	hypothetical protein FLJ10392
15	128888	RC_Z40388_s	A1760853	Hs.241558	ariadne (Drosophila) homolog 2
	125310	RC_Z40646	R59161	Hs.124953	ESTs
	125315	RC_Z41697	R38110	Hs.106296	ESTs
	125317	RC_Z98349	Z98348	Hs.112461	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	135086	RC_Z98394_s	AA081258	Hs.132390	zinc finger protein 36 (KIX 18)
20	104786	RC_AA027168	AA027167	Hs.10031	KIAA0955 protein
	123837	D58024_s	AA370362	Hs.57858	EGF-TM7-tatrophilin-related protein
	120456	RC_AA251113	AA488750	Hs.88414	BTB and CNC homology 1, basic leucine zipper transcription factor 2
	132459	RC_AA347573	AL120071	Hs.48998	fibronectin leucine rich transmembrane protein 2
	101545	M31210	BE246154	Hs.154210	endothelial differentiation, sphingolipid G-protein-coupled receptor, 1
25	133505	C01527	A1630124	Hs.324504	Homo sapiens mRNA; cDNA DKFZp586J0720 (from clone DKFZp586J0720)
	132360	RC_N62948_s	AW893660	Hs.46440	solute carrier family 21 (organic anion transporter), member 3
	132738	RC_W42674	AK000738	Hs.264636	hypothetical protein FLJ20731
	119586	RC_W43000_s	AF088033	Hs.159225	ESTs
	129914	RC_N31750_s	NM_012421	Hs.13321	rearranged L-myc fusion sequence
30	130839	AF009301	AB011169	Hs.20141	similar to S. cerevisiae SSM4
	132813	L37347	BE313625	Hs.57435	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2
	134342	M99564	NM_000275	Hs.82027	oculocutaneous albinism II (pink-eye dilution (murine) homolog)
	131878	RC_AA430673	AA083764	Hs.6101	hypothetical protein MGC3178
	105426	RC_AA251297	W20027	Hs.23439	ESTs
35	132968	RC_AA620722	AF234532	Hs.61638	myosin X
	132173	RC_W46577_s	X89426	Hs.41716	endothelial cell-specific molecule 1
	113932	RC_W81237	AA256444	Hs.126485	hypothetical protein FLJ12604; KIAA1692 protein
	114452	RC_AA020825	A1369275	Hs.243010	Homo sapiens cDNA FLJ14445 f1s, clone HEMBB1001294, highly similar to GTP-BINDING
		PROTEIN TC10			
40	115243	RC_AA278766	AA806600	Hs.116665	KIAA1842 protein
	134403	RC_H93708_s	AA334551	Hs.82767	sperm specific antigen 2
	129647	RC_N49394	AB018259	Hs.118140	KIAA0716 gene product
	111428	RC_H56559_s	AL031428	Hs.174174	KIAA0601 protein
	115967	RC_AA446887	A1745379	Hs.42911	ESTs
45	120726	RC_AA293656	AA293655	Hs.97293	ESTs
	114985	RC_AA251152	AA769266	Hs.193657	ESTs
	303876	RC_AA233334_s	U64820	Hs.66521	Machado-Joseph disease (spinocerebellar ataxia 3, olivopontocerebellar ataxia 3, autosomal dominant, ataxin 3)
	311483	RC_Z39920	R55344	Hs.22142	cytochrome b5 reductase b5R.2
	120302	RC_AA192173	AA837098	Hs.269933	ESTs
50	133071	RC_AA455044	BE384932	Hs.64313	ESTs, Weakly similar to AF257182 1 G-protein-coupled receptor 48 [H.sapiens]
	121032	RC_AA398504	AA393037	Hs.181798	ESTs
	128829	U41813	AF010258	Hs.127428	homeo box A9
	120245	RC_AA166965	AW959615	Hs.111045	ESTs
	120985	RC_AA398222	A1219896	Hs.97592	ESTs
55	114184	RC_Z39095	R56434	Hs.21062	ESTs
	447503	RC_AA284744_f	AA115496	Hs.336898	Homo sapiens, Similar to RIKEN cDNA 1810038N03 gene, clone MGC:9890, mRNA, complete cds
	132837	RC_AA428201	AA370362	Hs.57958	EGF-TM7-tatrophilin-related protein
	121034	RC_AA398507	AL389951	Hs.271623	nucleoporin 50kD
	119718	RC_W69216	W69216	Hs.92848	ESTs
60	120455	RC_AA251083	AA251720	Hs.104347	ESTs, Weakly similar to ALUC_HUMAN !!! ALU CLASS C WARNING ENTRY !!! [H.sapiens]
	125280	RC_W93659	A1123705	Hs.106932	ESTs
	132155	RC_AA227903	AK001607	Hs.41127	hypothetical protein FLJ13220
	120809	RC_AA283902	AW978721	Hs.266076	ESTs, Weakly similar to A46010 X-linked retinopathy protein [H.sapiens]
	121278	RC_AA401631	AA037121	Hs.98518	Homo sapiens cDNA FLJ11490 f1s, clone HEMBA1001918
65	109023	RC_AA157293	AA157293	Hs.72168	ESTs
	129815	RC_D60208_f	BE565817	Hs.26498	hypothetical protein FLJ21657
	108081	RC_AA043979	AA043979	Hs.62651	EST
	113287	RC_T68847	T68847	Hs.194040	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	114082	RC_Z38239	AK001612	Hs.26962	Homo sapiens cDNA FLJ10750 f1s, clone NT2RP3001929
70	116334	RC_AA491457	AL038450	Hs.48948	ESTs
	131486	RC_Z40071_s	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	107860	RC_AA024961	AA024961	Hs.50730	ESTs
	131263	RC_AA443826	ALU077002	Hs.24950	regulator of G-protein signalling 5
	132207	RC_AA443294	BE206839	Hs.42287	E2F transcription factor 6
75	129183	RC_AA155743	BE561824	Hs.273369	uncharacterized hematopoietic stem/progenitor cells protein MDS027
	408431	RC_T23708	A1338631	Hs.43266	Homo sapiens cDNA: FLJ22536 f1s, clone HRC13155
	120575	RC_AA280934	AW978022	Hs.238911	hypothetical protein DKFZp762E1511; KIAA1816 protein

	132121	RC_AA443284_s	NM_004529	Hs.404	myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 3
	117657	RC_N39074	N39074	Hs.44933	ESTs
	134922	RC_W04507_s	A1718295	Hs.91161	prefoldin 4
5	118523	RC_R41828_s	Y07759	Hs.170157	myosin VA (heavy polypeptide 12, myosin)
	118645	RC_H64973	AA649530		gbns44105.s1 NCL_CGAP_A1v1 Homo sapiens cDNA clone, mRNA sequence
	115291	RC_AA279943	BE545072	Hs.122579	hypothetical protein FLJ10461
	120326	RC_AA196300	AA196300	Hs.21145	hypothetical protein RG083M05.2
	130174	M29550	M29551	Hs.151531	protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform (calcineurin A beta)
10	129131	RC_AA436489	AB026436	Hs.177534	dual specificity phosphatase 10
	129668	RC_AA287032	AW172431	Hs.13012	ESTs
	118681	RC_N70777	AL137554	Hs.49927	protein kinase NYD-SP15
	129829	RC_AA496921	AF010258	Hs.127428	homeo box A9
	115985	RC_AA447709	AA447709	Hs.268115	ESTs, Weakly similar to T08599 probable transcription factor CA150 [H.sapiens]
15	134637	RC_AA368856_s	U87309	Hs.180941	vacuolar protein sorting 41 (yeast homolog)
	132714	RC_AA252598	W39388	Hs.55338	Homo sapiens, clone MGC:17421, mRNA, complete cds
	129771	RC_H73237	AL096748	Hs.102708	DKFZP434A043 protein
	123360	RC_AA504784	AA532718	Hs.178604	ESTs
	132902	RC_AA490969	A1938442	Hs.59838	hypothetical protein FLJ10808
20	113716	RC_T97750	AA001356	Hs.18159	ESTs
	113825	RC_W48860	AW014486	Hs.22509	ESTs
	130367	RC_Z38501	AL135301	Hs.8768	hypothetical protein FLJ10849
	120541	RC_AA278298	W07318	Hs.240	M-phase phosphoprotein 1
	116727	RC_F13684	R76472	Hs.65646	ESTs
25	118219	RC_N62231	AA862391	Hs.48494	ESTs, Moderately similar to A46010 X-linked retinopathy protein [H.sapiens]
	119767	RC_W72562	W72562	Hs.58119	ESTs
	128917	RC_AA481252	A1365215	Hs.206097	oncogene TC21
	451553	RC_AA020928	AA018454	Hs.269211	ESTs
	132716	RC_AA251288	BE379595	Hs.283738	casein kinase 1, alpha 1
	118525	RC_N67861	N67861	Hs.49390	ESTs
30	114618	RC_AA084162	AW979261	Hs.291993	ESTs
	119743	RC_W70242	AA947552	Hs.58086	ESTs
	108154	RC_AA425151_s	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain-binding protein
	122798	RC_AA460324	AW386286	Hs.145696	splicing factor (CC1.3)
35	133746	U44378	AW410035	Hs.75862	MAD (mothers against decapentaplegic, Drosophila) homolog 4
	119822	RC_W74471	AF086409	Hs.301327	ESTs
	122186	RC_AA435842	AA398811	Hs.104873	ESTs
	114941	RC_AA243017	AA236512	Hs.87331	ESTs
	118053	RC_N53367	N53391	Hs.47629	ESTs
40	123234	RC_AA490227	NM_001938	Hs.16697	down-regulator of transcription 1, TBP-binding (negative cofactor 2)
	129280	M63154	M63154	Hs.110014	gastric intrinsic factor (vitamin B synthesis)
	118995	RC_N94591	N94591	Hs.323056	ESTs
	116750	RC_H05960	AA760689	Hs.92418	ESTs
	129026	M98833	AL120297	Hs.108043	Friend leukemia virus integration 1
45	105127	RC_AA158132	AA045648	Hs.301957	nucleic acid diphosphate linked moiety X)-type motif 5
	114513	RC_AA044825	AA044873	Hs.103446	ESTs
	411856	RC_T35697	H67899	Hs.4190	Homo sapiens cDNA: FLJ23269 fis, clone COL09533
	132036	W01568	AL157433	Hs.37706	hypothetical protein DKFZp434E2220
	130091	RC_W88999	W88999		gb:zh70h03.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone 3', mRNA sequence
50	414108	U09564	A1267592	Hs.75761	SFRS protein kinase 1
	119881	RC_W81456	W81486	Hs.58648	ESTs
	117770	RC_N47953	AW957372	Hs.46791	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
	119850	RC_W80447	A1247568	Hs.58452	ESTs
55	115439	RC_AA284561	A1567972	Hs.193090	ESTs, Highly similar to AF161437 1 HSPC319 [H.sapiens]
	123107	RC_AA486071	AA225048	Hs.104207	ESTs
	406698	M24364	X03068	Hs.73931	major histocompatibility complex, class II, DQ beta 1
	112131	RC_AA400780	AA814948	Hs.96343	ESTs, Weakly similar to ALUC_HUMAN [!!!] ALU CLASS C WARNING ENTRY [!!!] [H.sapiens]
	132074	AB002366	AA478486	Hs.3852	KIAA0368 protein
60	413670	AB000115	AB000115	Hs.75470	hypothetical protein, expressed in osteoblast
	125277	RC_W93227	W93227	Hs.103245	EST
	114056	RC_AA186324	AA186175	Hs.82506	KIAA1254 protein
	121153	RC_AA399640	AA399640	Hs.97694	ESTs
	121609	RC_AA416867	AA416867	Hs.98185	EST
65	120661	RC_AA287556	AA287556	Hs.263412	ESTs, Weakly similar to ALUB_HUMAN [!!!] ALU CLASS B WARNING ENTRY [!!!] [H.sapiens]
	120850	RC_AA349647	AA349647	Hs.96927	Homo sapiens cDNA FLJ12573 fis, clone NT2RM4000979
	124947	RC_T03170	T03170	Hs.100165	ESTs
	130529	RC_AA280886	AA178953		gb:zp39e03.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 3' similar to contains Alu
70	117683	RC_N40180	N40180		repetitive element, mRNA sequence
	120745	RC_AA302809	AA302809		gb:yy44d02.s1 Soares_multiple_sclerosis_2NbhMSP Homo sapiens cDNA clone
	120936	RC_AA385934	AA385934	Hs.97184	gb:EST10426 Adipose tissue, white 1 Homo sapiens cDNA 3' end, mRNA sequence.
	112597	RC_R78376	R78376	Hs.29733	EST, Highly similar to (define not available 7499603) [C.elegans]
	120183	RC_Z40174	AW082866	Hs.65882	EST
75	120644	RC_AA287038	A1869129	Hs.96616	ESTs

	119023 RC_N98488	N98488	gbzdb2h01.s1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA clone
	IMAGE:310129.3'	mRNA sequence.	
	107582 RC_AA002147	AA002147	Hs.59952 EST
5	118249 RC_N62580	N62580	Hs.322925 EST, Weakly similar to putative p150 [H.sapiens]
	115022 RC_AA252029	AA252029	Hs.87935 ESTs
	117710 RC_N45198	N45198	Hs.47248 ESTs, Highly similar to similar to Cdc14B1 phosphatase [H.sapiens]
	115341 RC_AA281452	AA281452	Hs.88840 EST, Weakly similar to granule cell marker protein [M.musculus]
	118896 RC_N90680	N90680	Hs.54642 methionine adenosyltransferase II, beta
10	121121 RC_AA393371	AA393371	Hs.189095 similar to SALL1 (sal (Drosophila)-like
	116329 RC_N63520	N63520	gb:yy62f01.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone IMAGE:278137
	3', mRNA sequence.		
	119496 RC_W35416	W35416	Hs.156861 ESTs, Moderately similar to A46010 X-linked retinopathy protein [H.sapiens]
	118111 RC_N55493	N55493	gb:yy50c02.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:246146.3',
	mRNA sequence.		
15	119062 RC_R16898	AW444881	Hs.77829 ESTs
	116710 RC_F10577_f	F10577	Hs.306088 v-crk avian sarcoma virus CT10 oncogene homolog
	119261 RC_T15956	T15956	Hs.65289 EST
	122723 RC_AA457380	AA457380	gb:aa86b10.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:838171.3'
	similar to contains L1.b3 L1	repetitive element, mRNA sequence.	
20	117732 RC_N46452	N46452	gb:yy76h09.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone
	IMAGE:279521.3'	similar to contains L1.12 L1	repetitive element, mRNA sequence.
	104787 RC_AA027317	AA027317	gb:ze97d11.s1 Soares_fetal_heart_NbH19W Homo sapiens cDNA clone IMAGE:366933.3'
	similar to contains Alu	repetitive element, mRNA sequence.	
25	100071 A28102	A28102	Human GABAA receptor alpha-3 subunit
	115819 RC_AA426573	AA426573	Hs.41135 endomucin-2
	130882 RC_Z40166_f	AA497044	Hs.20887 hypothetical protein FLJ10392
	125225 RC_W76540	W74169	Hs.16492 DKFZP564G2022 protein
	106339 RC_AA070801	AW151340	Hs.51615 ESTs, Weakly similar to ALU7_HUMAN ALU SUBFAMILY SQ SEQUENCE CONTAMINATION
	WARNING ENTRY [H.sapiens]		
30	100338 D63483	D66864	Hs.57735 acetyl LDL receptor, SREC
	121636 RC_AA417027	AA379203	Hs.306654 Homo sapiens cDNA FLJ13574 fts, clone PLACE1008625
	103875 RC_AA418387	T26379	Hs.48802 Homo sapiens clone 23632 mRNA sequence
	118716 RC_N73460	AI658908	Hs.118722 fucosyltransferase 8 (alpha (1,6) fucosyltransferase)
	119763 RC_W72450	R54146	Hs.10450 Homo sapiens cDNA: FLJ22063 fts, clone HEP10326
35	121917 RC_AA428218	AA406397	Hs.98038 ESTs
	132806 M91488	AI699432	Hs.278619 hypothetical protein FLJ10099
	130949 Y10659	AV656840	Hs.285115 interleukin 13 receptor, alpha 1
	108806 RC_AA129933	AF070578	Hs.71168 Homo sapiens clone 24674 mRNA sequence
	133276 RC_AA490478	AW978439	Hs.69504 ESTs
40	134760 RC_H16758	NM_000121	Hs.89548 erythropoietin receptor
	132867 AA121287	AF226667	Hs.58553 CTP synthase II
	132051 AA091284	AA393968	Hs.180145 HSPC030 protein
	114208 RC_Z39301	AL049466	Hs.7859 ESTs
	104094 AA418187	AA418187	Hs.330515 ESTs
45	128718 AA426361	NM_002959	Hs.281706 sortilin 1
	302032 RC_N20407	NM_001992	Hs.128087 coagulation factor II (thrombin) receptor
	115501 RC_AA291553	AA291553	Hs.190086 ESTs
	101997 U01160	AU076536	Hs.50984 sarcoma amplified sequence
50	103708 AA037206	AA430591	Hs.72071 hypothetical protein FLJ20038
	101899 S59184	S59184	Hs.79350 RYK receptor-like tyrosine kinase
	115839 RC_AA429038	BE300266	Hs.28935 transducin-like enhancer of split 1, homolog of Drosophila E(sp1)
	409459 D50678	D66407	Hs.54481 low density lipoprotein receptor-related protein 8, apolipoprotein e receptor
	103563 Z22534	L02911	Hs.150402 Activin A receptor, type I (ACVR1) (ALK-2)
55	123233 RC_AA490225	AW974175	Hs.188751 ESTs, Weakly similar to MAPB_HUMAN MICROTUBULE-ASSOCIATED PROTEIN 18
	[H.sapiens]		
	121305 RC_AA402468	AA402468	Hs.291557 ESTs
	114798 RC_AA159181	AA159181	Hs.54900 serologically defined colon cancer antigen 1
	133145 RC_AA196549	H94227	Hs.6592 Homo sapiens, clone IMAGE:2961368, mRNA, partial cds
60	131567 RC_AA291015_s	AF015592	Hs.28853 CDC7 (cell division cycle 7, S. cerevisiae, homolog)-like 1
	112300 RC_R54554	H24334	Hs.26125 ESTs
	129507 RC_AA192099	AJ236885	Hs.112180 zinc finger protein 148 (pH2-52)
	121033 RC_AA398505	AA398505	Hs.97360 ESTs
	121151 RC_AA399636	AA399636	Hs.143629 ESTs
	121402 RC_AA406063	AA406063	Hs.98003 ESTs
65	123203 RC_AA489671	AA352335	Hs.65641 hypothetical protein FLJ20073
	132271 RC_AA236466	AB030034	Hs.115175 sterile-alpha motif and leucine zipper containing kinase AZK
	125197 RC_W69106	AF086270	Hs.278554 heterochromatin-like protein 1
	114935 RC_AA242809	H23329	Hs.290880 ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION
	WARNING ENTRY [H.sapiens]		
70	125279 RC_W93640	AW401809	Hs.4779 KIAA1150 protein
	108778 RC_AA128548	AF133123	Hs.90847 general transcription factor IIIc, polypeptide 3 (102kD)
	108087 RC_AA045709	AA045709	Hs.40545 ESTs
	132466 RC_N66810_s	AI597655	Hs.49265 ESTs
	133328 R36553	AW452738	Hs.265327 hypothetical protein DKFZp7611141
75	124057 RC_F13604	AA902384	Hs.73853 bone morphogenetic protein 2
	124800 RC_R45115	AW864086	Hs.138617 thyroid hormone receptor interactor 12

	121029	RC_AA398482	AA398482	Hs.97641	EST
	120663	RC_AA287627	AA827798	Hs.105089	ESTs
	102133	U15173	AU076845	Hs.155596	BCL2/adenovirus E1B 19kD-interacting protein 2
5	108246	RC_AA062855	AA23132	Hs.146343	ESTs
	125226	RC_W78134	AA782536	Hs.122647	N-myristoyltransferase 2
	120260	RC_AA171739	AK000001	Hs.101590	hypothetical protein
	124906	RC_R87647	H75984	Hs.107815	ESTs
	109406	RC_AA226877	AA199883	Hs.87624	ESTs
	109271	RC_AA195668	AW137422	Hs.86022	ESTs
10	125052	RC_T80174_s	T85104	Hs.222779	ESTs, Moderately similar to similar to NEDD-4 [H.sapiens]
	109101	RC_AA167708	AW608930	Hs.52184	hypothetical protein FLJ20618
	115241	RC_AA278723	AA648278	Hs.193859	ESTs
	117163	RC_H97909	N36861	Hs.42344	ESTs
	113530	RC_T90313	T90313	Hs.16732	ESTs
15	120375	RC_AA227260	AF028706	Hs.111227	Zic family member 3 (odd-paired Drosophila homolog, heterotaxy 1)
	129435	AA314256	AF151852	Hs.111449	CGI-94 protein
	114864	RC_AA235256	AA135332	Hs.71608	ESTs
	103988	AA314389	AA314389	Hs.42500	ADP-ribosylation factor-like 5
20	131006	RC_AA242763	AF064104	Hs.22116	CDC14 (cell division cycle 14, S. cerevisiae) homolog B
	106781	RC_AA478474	AA330310	Hs.24181	ESTs
	106141	RC_AA424558	AF031463	Hs.9302	phosphatase-like
	116213	RC_AA476738	AA292105	Hs.326740	hypothetical protein MGC10947
	135266	AB002326	R41179	Hs.97393	KIAA0328 protein
25	135058	RC_AA430152	AI379720	Hs.93814	hypothetical protein
	119908	RC_W85844	AA524470	Hs.58753	ESTs
	103695	AA018758	AW207152	Hs.186600	ESTs
	103978	AA307443	NM_016940	Hs.34136	chromosome 21 open reading frame 6
	109485	RC_AA233472	BE619092	Hs.28465	Homo sapiens cDNA: FLJ21869 fis, clone HEP02442
30	129574	AA458603	AA026815	Hs.11463	UMP-CMP kinase
	115347	RC_AA281528	AA356792	Hs.334824	hypothetical protein FLJ14825
	120765	RC_AA338735	AW961026	Hs.96752	ESTs, Weakly similar to ALU8_HUMAN ALU SUBFAMILY SX SEQUENCE CONTAMINATION
	WARNING ENTRY [H.sapiens]				
	121059	RC_AA398628	AA393283		gb:z174e03.1 Soares_testis_NHT Homo sapiens cDNA clone 5', mRNA sequence
35	131887	AA046548	W17064	Hs.332848	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1
	112064	RC_R43812	AL049390	Hs.22689	Homo sapiens mRNA; cDNA DKFZp586O1318 (from clone DKFZp586O1318)
	115606	RC_AA400465	AI025829	Hs.86320	ESTs
	131750	RC_H94855_s	NM_004349	Hs.31551	core-binding factor, runt domain, alpha subunit 2; translocated to, 1; cyclin D-related
40	102123	U14518	NM_001809	Hs.1594	centromere protein A (17kD)
	129847	RC_W46767	N64025	Hs.296178	hypothetical protein FLJ22637
	133809	RC_AA235275	AV649326	Hs.76359	catalase
	132210	RC_N51499_s	NM_007203	Hs.42322	A kinase (PRKA) anchor protein 2
	122356	RC_AA443794	AA443794	Hs.98390	ESTs
	114958	RC_AA243708	N20912	Hs.42369	ESTs
45	103951	AA287840	AL353944	Hs.50115	Homo sapiens mRNA; cDNA DKFZp761J1112 (from clone DKFZp761J1112)
	134703	RC_AA280704	AF117065	Hs.88764	male-specific lethal-3 (Drosophila)-like 1
	128727	AA287864	AI223335	Hs.50651	Janus kinase 1 (a protein tyrosine kinase)
	105743	RC_AA293300_s	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4B
50	103744	AA076003	AA079267		gb:zm97e10.s1 Stratagene colon HT29 (937221) Homo sapiens cDNA clone 3', mRNA sequence
	114348	N80402	AL050321	Hs.301532	CRP2 binding protein
	114009	RC_W90067	AI248544	Hs.103000	KIAA0831 protein
55	134704	RC_AA280849	AA837124	Hs.88780	ESTs
	128629	AA399187	AL096748	Hs.102708	DKFZP434A043 protein
	104410	H85925	AI807519	Hs.104520	Homo sapiens cDNA FLJ13694 fis, clone PLACE2000115
	110200	RC_H21075	H21075	Hs.31802	ESTs, Highly similar to A59266 unconventional myosin-15 [H.sapiens]
	124483	RC_N53976	AI821780	Hs.179864	ESTs
60	101391	M14648	NM_002210	Hs.295726	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)
	109657	RC_F04826	R60900	Hs.26814	ESTs
	117140	RC_H96813	H96813	Hs.42241	ESTs
	132937	RC_AA233706_f	AW952912	Hs.300383	hypothetical protein MGC3032
	129799	R36410	AW967473	Hs.239114	mannosidase, alpha, class 1A, member 2
65	105077	RC_AA142919	W55946	Hs.234863	Homo sapiens cDNA FLJ12082 fis, clone HEMBB1002492
	100850	RC_N58561_s	AA836472	Hs.297939	cathepsin B
	131043	RC_AA490925	AF084535	Hs.22464	epilepsy, progressive myoclonus type 2, Lafora disease (laforin)
	118417	RC_N66048_f	AF080229		gb:human endogenous retrovirus K clone 10.1 polymerase mRNA, partial cds
	129254	RC_AA243695	AA252468	Hs.1098	DKFZp434J1813 protein
70	119149	RC_R58910	BE304701	Hs.65732	ESTs
	133996	AA091367	AA380267	Hs.78277	DKFZP434F2021 protein
	110223	RC_H23747	H19836	Hs.31697	ESTs
	117626	RC_N36090	AK001757	Hs.281348	hypothetical protein FLJ10895
	135288	RC_AA424469_s	AW023482	Hs.97849	ESTs
75	122967	RC_AA478521	AA806187	Hs.289101	glucose regulated protein, 58kD
	131236	AA282640	AF043117	Hs.24594	ubiquitination factor E4B (homologous to yeast UFD2)
	128568	AA463380	H12912	Hs.274691	adenylylase kinase 3

	112888	RC_T03872	AW195317	Hs.107716	hypothetical protein FLJ22344
	115192	RC_AA261920	AA741024	Hs.88378	ESTs
	118688	RC_N71484	AK000708	Hs.169764	hypothetical protein FLJ20701
	122264	RC_AA436837	AA436837		gbczv57g07.s1 Soares_testis_NHT Homo sapiens cDNA clone 3', mRNA sequence
5	128981	AA135452	AA927177	Hs.86041	CGG triplet repeat binding protein 1
	131042	RC_R42457	AI826288	Hs.171637	hypothetical protein MGC2628
	103704	AA028171	AA028171	Hs.151258	hypothetical protein FLJ21062
	121341	AA233107	AF035528	Hs.153863	MAD (mothers against decapentaplegic, Drosophila) homolog 6
	106593	RC_AA456826	AW296451	Hs.24605	ESTs
10	115195	RC_AA262156	AW968619	Hs.155849	ESTs
	115425	RC_AA284071	AA811895	Hs.180680	ESTs, Weakly similar to I54374 gene NF2 protein [H.sapiens]
	117258	RC_N21299	AF086041	Hs.42975	ESTs
	120209	RC_Z40892	F02951		gbtHSC1HB082 normalized infant brain cDNA Homo sapiens cDNA clone c-1hb08 3', mRNA sequence
15	134082	L16991	L16991	Hs.79006	deoxythymidylate kinase (thymidylate kinase)
	104774	RC_AA026066	AW959755	Hs.288896	Homo sapiens cDNA FLJ12977 fis, clone NT2RP2006261
	115625	RC_AA401630	AA059459	Hs.62592	ESTs
	104469	N28707	N28707	Hs.154304	Homo sapiens chromosome 19, BAC 282485 (CIT-B-344H19)
	107401	W20054	N91453	Hs.102987	ESTs
20	111686	RC_R21510	R22039	Hs.23217	ESTs
	115300	RC_AA280026	AA280095	Hs.88689	ESTs
	115378	RC_AA282292	AA282292	Hs.279841	hypothetical protein FLJ10335
	132224	RC_H97819	N41549	Hs.285410	ESTs
	113791	M95767	AI269096	Hs.135578	chitinase, di-N-acetyl-
25	129144	AA004987	AL137275	Hs.20137	hypothetical protein DKFZp434P0116
	104448	L44574	NM_007331	Hs.110457	Wolff-Hirschhorn syndrome candidate 1
	132084	RC_T26981_s	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)
	111831	RC_R36083	R36095	Hs.268695	ESTs
	114765	RC_AA252163	AA463550	Hs.337532	ESTs, Weakly similar to A47582 B-cell growth factor precursor [H.sapiens]
30	115029	RC_AA252219	AL137939	Hs.40096	ESTs
	100457	H81492	BE246400	Hs.285176	acetyl-Coenzyme A transporter
	104536	R24011	R24024	Hs.158101	Homo sapiens cDNA FLJ14673 fis, clone NT2RP2003714, moderately similar to ZINC FINGER PROTEIN 91
	116167	RC_AA461562	AI091731	Hs.87293	hypothetical protein FLJ20045
35	103889	AA236771	R85350	Hs.101368	ESTs
	131978	RC_H48459_s	AA355925	Hs.36232	KIAA0186 gene product
	118843	RC_N80181	N80181	Hs.221498	ESTs
	120837	RC_W93092	BE149656	Hs.306621	Homo sapiens cDNA FLJ11963 fis, clone HEMBB1001051
	133647	D21852	NM_015361	Hs.268053	KIAA0029 protein
40	129521	U41815	AF071076	Hs.112255	nucleoporin 98kD
	103746	AA081876	AA075000		gb:zm83c07.s1 Stratagene ovarian cancer (937219) Homo sapiens cDNA clone 3', mRNA sequence
	132019	RC_AA134955_J	H56995	Hs.37372	Homo sapiens DNA binding peptide mRNA, partial cds
	132310	RC_AA284107	AA173223	Hs.289044	Homo sapiens cDNA FLJ12048 fis, clone HEMBB1001990
45	117367	RC_N24954	AI041793	Hs.42502	ESTs
	103743	AA075998	AA075998		gb:zm89b09.r1 Stratagene ovarian cancer (937219) Homo sapiens cDNA clone 5' similar to gb:zm79b10.s1 NCL_CGAP_GCB1 Homo sapiens cDNA clone 3' similar to gb:M34539 FK506-BINDING PROTEIN (HUMAN); mRNA sequence
	103761	AA085138	AA765163		gb:nz79b10.s1 NCL_CGAP_GCB1 Homo sapiens cDNA clone 3' similar to gb:M34539 FK506-BINDING PROTEIN (HUMAN); mRNA sequence
50	130237	L39060	AA913909	Hs.153088	TATA box binding protein (TBP)-associated factor, RNA polymerase I, A, 48kD
	128752	RC_N72879	AA504428	Hs.10487	Homo sapiens, clone IMAGE:3954132, mRNA, partial cds
	135162	AA045930	AI187925	Hs.95667	F-box protein 30
	131386	AA086412	BE219898	Hs.173135	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2
	129021	RC_AA599244	AL044675	Hs.173081	KIAA0530 protein
55	424274	AA293634	W73933	Hs.283738	casein kinase 1, alpha 1
	129913	H06583	NM_001310	Hs.13313	cAMP responsive element binding protein-like 2
	131888	U79298	AW294659	Hs.34054	Homo sapiens cDNA: FLJ22488 fis, clone HRC10948, highly similar to HSU79298 Human clone
	23803	mRNA			
	118612	RC_N69466	AB037788	Hs.224961	cleavage and polyadenylation specific factor 2, 100kD subunit
60	322026	AA203138	AW024973	Hs.283675	NPD009 protein
	110892	RC_N38882	AL035301	Hs.97375	H.sapiens gene from PAC 106H8
	111429	RC_R01245	AI038052	Hs.19162	ESTs, Weakly similar to I54374 gene NF2 protein [H.sapiens]
	113334	RC_T76962	AW974666	Hs.283024	ESTs
	104091	AA417310	BE465093	Hs.106101	hypothetical protein FLJ22557
65	105246	RC_AA226879	AA226879		gb:zt19c09.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:663856 3' similar to contains Alu repetitive element, mRNA sequence.
	113300	RC_T67448	T67448	Hs.13101	ESTs
	117147	RC_H97225_s	AW901347	Hs.38592	hypothetical protein FLJ23342
	121349	RC_AA405205	AA405205	Hs.97960	ESTs, Weakly similar to T51146 ring-box protein 1 [H.sapiens]
70	100294	D49396	AA331881	Hs.75454	peroxiredoxin 3
	133999	M28213	AA535244	Hs.78305	RAB2, member RAS oncogene family
	133259	AA278548	BE379646	Hs.6904	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 2004403
	129423	AA371418	AA204686	Hs.234149	hypothetical protein FLJ20647
	131098	RC_AA459668	U66669	Hs.236642	3-hydroxyisobutyryl-Coenzyme A hydrolase
75	135272	AA399391	AI828337	Hs.97591	ESTs
	129155	AA046865	AI952677	Hs.108972	Homo sapiens mRNA; cDNA DKFZp434P228 (from clone DKFZp434P228)

	311291	AA056319	AA782601	Hs.319817	ESTs
	120750	RC_AA310499	AI191410	Hs.96693	ESTs, Moderately similar to 2109260A B cell growth factor [H.sapiens]
	101002	J04058	AV655843	Hs.189919	electron-transfer-flavoprotein, alpha polypeptide (glutamic aciduria II)
5	133012	AA099241	AA847843	Hs.62711	Homo sapiens, clone IMAGE:3351295, mRNA
	103879	AA228148_s	BE543269	Hs.50252	mitochondrial ribosomal protein L32
	131281	RC_AA443212	AA251716	Hs.25227	ESTs
	15109	RC_AA256383	AJ249977	Hs.88049	protein kinase, AMP-activated, gamma 3 non-catalytic subunit
	18502	RC_N67317	AL157488	Hs.50150	Homo sapiens mRNA; cDNA DKFZp564B182 (from clone DKFZp564B182)
10	134100	L07540	AA60085	Hs.171075	replication factor C (activator 1) 5 (36.5KD)
	131869	AA484944	AW968547	Hs.33540	ESTs, Weakly similar to cJ309K20.4 [H.sapiens]
	115398	RC_AA282985	AA810854	Hs.89081	ESTs
	103860	AA203742	AW976877	Hs.38057	ESTs
	135089	N75811_s	AI918035	Hs.301198	roundabout (axon guidance receptor, Drosophila) homolog 1
	128938	U79300	AW003668	Hs.135587	Human clone 23629 mRNA sequence
15	107508	W90095	N74925	Hs.38761	Homo sapiens cDNA: FLJ21564 fls, clone COL06452
	103685	AA005190	AA158008	Hs.292444	ESTs
	125170	AA203147	AL020996	Hs.8518	selenoprotein N
	129179	RC_AA504125_s	AW969025	Hs.109154	ESTs
	116262	AA477046	AI936442	Hs.59838	hypothetical protein FLJ10808
20	123009	RC_AA479949	AA535244	Hs.78305	RAB2, member RAS oncogene family
	131004	D29833	D29833	Hs.2207	salivary proline-rich protein
	103317	X83441	X83441	Hs.166091	ligase IV, DNA, ATP-dependent
	132814	RC_C15251_f	D60730	Hs.57471	ESTs
25	103992	U77718	BE018142	Hs.300954	Huntingtin interacting protein K
	108258	X59710	AL044818	Hs.84928	nuclear transcription factor Y, beta
	110754	RC_N20814	AW302200	Hs.6336	KIAA0672 gene product
	132727	AA136382_s	N27495	Hs.5565	hypothetical protein FLJ22626
	100341	D63506	AF032922	Hs.8813	synaptobrevin binding protein 3
30	134684	AA256106	AA256106	Hs.87507	ESTs
	103826	AA165564	AW162998	Hs.24684	KIAA1376 protein
	111678	RC_R20628	R38487	Hs.169927	ESTs
	101341	L76159	NM_004477	Hs.203772	FSHD region gene 1
	115455	RC_AA285068	AA876002	Hs.120551	toil-like receptor 10
	111192	RC_AA477748	AW021968	Hs.109438	Homo sapiens clone 24775 mRNA sequence
35	129385	RC_AA235604	AA172106	Hs.110950	Rag C protein
	125050	RC_T79951	AW970209	Hs.111805	ESTs
	122105	RC_AA432278	AW241685	Hs.98699	ESTs
	121324	RC_AA404229	AA404229	Hs.97842	EST
40	120938	RC_AA386260	AA386260	Hs.104632	EST
	115001	RC_AA251376	AA251376		gbzsa10a06.s1 NCLCGAP_GCB1 Homo sapiens cDNA clone IMAGE:684754 3', mRNA sequence.
	124799	RC_R45088	R45088		gbzyg38g04.s1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:34896 3', mRNA sequence.
45	122724	RC_AA457395	AA457395	Hs.99457	ESTs
	117791	RC_N48325	N48325	Hs.93956	EST
	121895	RC_AA427396	AA427396		gbzaw33a02.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:771050 3'
	108244	RC_AA082839	AA082839		similar to contains Alu repetitive element; contains MER12.12 MER12 repetitive element ; mRNA sequence.
50	117852	RC_N49408	AW877787	Hs.136102	KIAA0853 protein
	109298	RC_AA205432	R77854	Hs.250693	Kruppel-related zinc finger protein
	122432	RC_AA447400	AA447400	Hs.187684	ESTs, Weakly similar to B34087 hypothetical protein [H.sapiens]
	124627	RC_N74625	N74625		gbzsa55c03.s1 Soares fetal liver spleen 1N1LS Homo sapiens cDNA clone IMAGE:296452 3'
55	115141	RC_AA258071	AA465131	Hs.64001	similar to gbz1M14338 VITAMIN K-DEPENDENT PROTEIN S PRECURSOR (HUMAN); contains OFR13 OFR repetitive element ; mRNA sequence.
	128636	U49065	U49065	Hs.102865	Homo sapiens clone 25218 mRNA sequence
	115373	RC_AA282197	AA664862	Hs.181022	interleukin 1 receptor-like 2
	114651	RC_AA101400	AA101400	Hs.189960	CGI-07 protein
60	132796	RC_AA180487	NM_006283	Hs.173159	ESTs
	103749	RC_N35583	AL135301	Hs.8768	transforming, acidic coiled-coil containing protein 1
	107328	T83444	AW959891	Hs.76591	hypothetical protein FLJ10849
	115349	RC_AA281563	AF121176	Hs.12797	KIAA0887 protein
	111490	RC_R06862	R06862		DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 16
65	103763	AA085354	AA085291		gbzyf11e09.s1 Soares fetal liver spleen 1N1LS Homo sapiens cDNA clone IMAGE:126568 3'
	118791	RC_N75520	N75520	Hs.261003	similar to contains L1 repetitive element ; mRNA sequence.
	116644	RC_F03032	F03032	Hs.290278	gbzbn01g06.s1 Stratagene colon HT29 (937221) Homo sapiens cDNA clone 3' similar to contains Alu repetitive element ; mRNA sequence
70	116823	RC_H56485	AW204742	Hs.143542	ESTs, Moderately similar to B34087 hypothetical protein [H.sapiens]
	108940	RC_AA148603	AA148603		ESTs, Weakly similar to B34087 hypothetical protein [H.sapiens]
	112218	RC_R50057	R50057	Hs.272251	ESTs, Highly similar to CSA_HUMAN COCKAYNE SYNDROME WD-REPEAT PROTEIN CSA [H.sapiens]
75	116557	RC_D20572_j	D20572	Hs.90171	gbzco09e04.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone IMAGE:567198 3', mRNA sequence.
	133649	U25849	U25849	Hs.75393	Homo sapiens mRNA; cDNA DKFZp586M1418 (from clone DKFZp586M1418)
	131745	RC_C20746	AI828559	Hs.31447	EST
					acid phosphatase 1, soluble
					ESTs, Moderately similar to A46010 X-linked retinopathy protein [H.sapiens]

	116801	RC_H43879	H43879		gbcyo69h09.s1 Soares breast 3NbhBst Homo sapiens cDNA clone IMAGE:183233 3', mRNA sequence.
	115006	RC_AA251548	AA251548	Hs.87886	EST
	123424	RC_AA598500	M29882	Hs.162614	ESTs
5	120831	RC_AA347919	AA347919	Hs.96889	EST
	103691	AA018298	AA018298	Hs.103332	ESTs
	121556	RC_AA412491	AF025771	Hs.50123	zinc finger protein 189
	111193	RC_N67946	N67946	Hs.117569	ESTs
	132061	RC_AA058946	AB020700	Hs.3830	KIAA0893 protein
10	134575	RC_AA194568_J	AA194568	Hs.85938	EST
	115050	RC_AA252794	AA252794	Hs.88009	ESTs
	420208	U31799	BE276055	Hs.95972	silver (mouse homolog) like
	133735	AC002045_xpt1	R66740	Hs.110613	KIAA0220 protein
	128546	Z21305	NM_003478	Hs.101299	culin 5
15	111946	RC_R40697	R40697	Hs.76666	C9orf10 protein
	124879	RC_R73588	R73588	Hs.101533	ESTs
	115683	AA410345	AF255910	Hs.54650	junctional adhesion molecule 2
	103692	AA018418	AW137912	Hs.227583	Homo sapiens chromosome X map Xp11.23 L-type calcium channel alpha-1 subunit
20	(CACNA1F) gene, complete cds; HSP27 pseudogene, complete sequence; and JM1 protein, JM2 protein, and Hb2E genes, complete cds				
	103767	AA089688	BE244667	Hs.296155	CGI-100 protein
	125266	W90022	Hs.186809		ESTs, Highly similar to LCT2_HUMAN LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2
	135235	AA435512	AW298244	Hs.293507	ESTs
	134497	RC_AA404494	BE258532	Hs.251871	CTP synthase
25	426754	RC_AA278529_J	NM_014264	Hs.172052	serine/threonine kinase 18
	412177	RC_AA342828_s	Z23091	Hs.73734	glycoprotein V (platelet)
	132000	RC_AA044644	AW247017	Hs.36978	melanoma antigen, family A, 3
	124738	RC_AA044644	T07568	Hs.137158	ESTs
	324000	RC_AA196729_J	AA604749	Hs.190213	ESTs
30	106896	RC_AA196729_J	AW073202	Hs.334825	Homo sapiens cDNA FLJ14752 fis, clone NT2RP3003071
	132000	RC_AA025858	AW247017	Hs.36978	melanoma antigen, family A, 3
	129577	RC_AA025858	N75346	Hs.82906	CDC20 (cell division cycle 20, S. cerevisiae, homolog)
	107091	RC_AA233519	AI949109	Hs.246885	hypothetical protein FLJ20783
	130296	RC_N52271	D31139	Hs.154103	LIM protein (similar to rat protein kinase C-binding enigma)
35	102855	RC_N68399	NM_003528	Hs.2178	H2B histone family, member Q
	113689	RC_AA098874	AB037850	Hs.16621	DKFZP434I116 protein
	100939	RC_AA279667_s	L04288	Hs.297939	cathepsin B
	130430	RC_H22556	W27893	Hs.150580	putative translation initiation factor
40	106734	RC_N45979_s	BE296690	Hs.288173	Homo sapiens cDNA: FLJ21747 fis, clone COLF5160, highly similar to AF182198 Homo sapiens
	intersectin 2 long isoform (ITSN2) mRNA				
	135148	RC_AA431288_s	AA306478	Hs.95327	CD3D antigen, delta polypeptide (TIT3 complex)
	134221	RC_AA609862	BE280456	Hs.80248	RNA-binding protein gene with multiple splicing
	105376	RC_N35583	AW994032	Hs.8768	hypothetical protein FLJ10849
45	124541	U77718	AF112222	Hs.44499	pinin, desmosome associated protein
	134546	AA203147	AL020996	Hs.8518	selenoprotein N
	134000	RC_W93092	AW175787	Hs.334841	selenium binding protein 1
	125656	RC_W93092	AW516428	Hs.78687	neutral sphingomyelinase (N-SMase) activation associated factor
	100939	RC_N58561_s	L04288	Hs.297939	cathepsin B
	125656	RC_W93092	AW516428	Hs.78687	neutral sphingomyelinase (N-SMase) activation associated factor
50	101779	RC_W69385_s	BE543412	Hs.250505	retinoic acid receptor, alpha
	332489	RC_R22947	R23053	NA	Hu01 Chip Redos
	133000	RC_N38959_f	AL042444	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast Sta20-related)
	125905	RC_N38959_f	AI678638	Hs.6456	chaperonin containing TCP1, subunit 2 (beta)
	129000	RC_H73050_s	AA744902	Hs.107767	hypothetical protein PRO1489
55	100920	RC_H73050_s	X54534	Hs.278994	Rhesus blood group, CcEe antigens

TABLE 1A

5 Table 1A shows the accession numbers for those pkeys lacking unigenelD's for Tables 1. The pkeys in Table 7 lacking unigenelD's are represented within Tables 1-6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10	Pkey:	Unique Eos probeset Identifier number	
	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
15	Pkey	CAT Number	Accession
	108469	116761_1	AA079487 AA128547 AA128291 AA079587 AA079600
	124106	125446_1	H12245 AA094769 R14576
	108501	13684_-12	AA083256
20	108562	36375_1	AA100796 AF020589 AA074629 AA075946 AA100849 AA085347 AA126309 AA079311 AA079323 AA085274
	125008	1802095_1	T91251 T84891 T85665
	125020	116017_1	T69981 T69924 AA078476
	125066	1814993_1	T86284 T81933
	116661	1532859_1	R81504 F04247
25	125104	413347_1	T95590 AA703278 H62764
	124575	1666649_1	N68168 N69188 N90450
	125263	1547_2	AA098878 W88942
	116845	393481_1	AA649530 AA659316 H64973
30	118417	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 A1636743 AW614951 BE467547 A1680833
			A1633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 A1970376 A1583718 A1672574
			N25695 AW665466 A1818326 AA126128 A1480345 AW013827 AA248638 A1214968 AA204735 AA207155 AA206262 AA204833
			AW003247 AW496808 A1080480 A1631703 A1651023 A1867418 AW818140 AA502500 A1206199 A1671282 A1352545 BE501030
			A1652535 BE465762 AA206331 AW451866 AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048
			AA703398 H92278 AW139734 H92683 U87589 U87595 H69001 U87594 BE466420 A1624817 BE466611 A1206344 AA574397
35			AA348354 A1493192
	118584	532052_1	AW136928 A1685655 BE218584 BE465078 N68963 AA975338 BE147199 N76377
	103743	112194_1	AA075998 AA075999 AA070986 AA070896 AA129207 AA078942 AA070783 AA078941
	103744	114161_1	AA079267 AA076003
	103746	113452_1	AA075000 AA081876
40	103761	114208_1	AA765163 AW298222 AA126126 AA085138 AA076068
	103763	48290_6	AA085291 AA085354
	120209	1531817_1	F02951 Z40892 F04711
	120284	158963_1	AA179656 AA182626 AA182603
	112540	1605263_1	R69751 R70467 H69771 H80879 H80878
45	111904	1719336_1	Z41572 R39330
	121059	273450_1	AA393283 AA398628
	121094	275729_1	AA402505 AA398900
	114106	1182096_1	AW602528 BE073859 Z38412
	130091	23961_-3	W88999
50	122264	296527_1	AA436837 AA442594
	108280	110682_1	AA065069 AA085108
	129961	1706092_1	R23053 R79884 R76271
	130529	158447_1	AA178953 AA192740
	108309	111495_1	AA069818 AA069971 AA069923 AA069908
55	107832	genbank_AA021473	AA021473
	123731	genbank_AA609839	AA609839
	116571	genbank_D45652	D45652
	132225	genbank_AA128980	AA128980
	125017	genbank_T68875	T68875
60	125083	genbank_T85352	T85352
	125064	genbank_T85373	T85373
	100964	entrez_J00212 J00212	
	125118	149288_1	R10608 T97620 AA576309
	102269	entrez_U30245 U30245	
65	125160	NOT_FOUND_entrez_W38240	W38240
	116801	genbank_H43879	H43879
	118111	genbank_N55493	N55493
	118129	genbank_N57493	N57493
	118329	genbank_N63520	N63520
70	118475	genbank_N66845	N66845
	111490	genbank_R06862	R06862
	111514	genbank_R07998	R07998
	104534	R22303_at	R22303
	120340	genbank_AA206828	AA206828

WO 02/079492

PCT/US02/04915

	120376	genbank_AA227469	AA227469
	104787	genbank_AA027317	AA027317
	120409	genbank_AA235050	AA235050
5	120745	genbank_AA302809	AA302809
	120809	genbank_AA346495	AA346495
	120839	genbank_AA348913	AA348913
	113702	genbank_T97307	T97307
	115001	genbank_AA251376	AA251376
	122562	genbank_AA452156	AA452156
10	122635	genbank_AA454085	AA454085
	108244	genbank_AA062839	AA062839
	108277	genbank_AA064859	AA064859
	122723	genbank_AA457380	AA457380
	124028	genbank_F04112	F04112
15	108403	genbank_AA075374	AA075374
	122860	genbank_AA464414	AA464414
	108427	genbank_AA076382	AA076382
	108439	genbank_AA078986	AA078986
	131353	231290_1	AW411259 H23555 AW015049 AJ684275 AW015886 AW068953 AW014085 AJ027260 R52686 AA918278 AJ129462
20	AA969360		
			N34859 AJ948418 AA534205 AA702483 AA705292
	106533	genbank_AA084415	AA084415
	117031	genbank_H88353	H88353
	124254	genbank_H69899	H69899
25	101447	entrez_M21305	M21305
	101458	entrez_M22092	M22092
	124577	genbank_N68300	N68300
	108940	genbank_AA148603	AA148603
	108941	genbank_AA148650	AA148650
30	124627	genbank_N74625	N74625
	124720	144582_1	R05283 R11056
	124793	genbank_R44519	R44519
	124799	genbank_R45088	R45088
	117683	genbank_N40180	N40180
35	117732	genbank_N46452	N46452
	124991	genbank_T50116	T50116
	119023	genbank_N98488	N98488
	119239	95573_2	T11483 T11472
	119558	NOT_FOUND_entrez_W38194	W38194
40	119654	genbank_W57759	W57759
	105246	genbank_AA226879	AA226879
	121350	genbank_AA405237	AA405237
	121558	genbank_AA412497	AA412497
	105985	genbank_AA406610	AA406610
45	100071	entrez_A28102A28102	
	114648	genbank_AA101056	AA101056
	121895	genbank_AA427396	AA427396
	100327	entrez_D55640D55640	
	123315	714071_1	AA496369 AA496646

TABLE 2:

5	Pkey:	Unique Eos probeset identifier number			
	Accession:	Accession number used for previous patent filings			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
10	Pkey	Accession	ExAccn	UnigeneID	UnigeneTitle
15	100420	100420	D86983	Hs.118893	Melanoma associated gene
	100484	100484	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene hom
	100991	100991	J03836	Hs.82085	serine (or cysteine) proteinase inhibitor
	101168	101168	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	101261	101261	D30857	Hs.82353	protein C receptor, endothelial (EPCR)
20	101447	101447	M21305		gbHuman alpha satellite and satellite 3
	101543	101543	M31166	Hs.2050	pentaxin-related gene, rapidly induced b
	101560	101560	AW958272	Hs.347328	intercellular adhesion molecule 2
	101714	101714	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
	101838	101838	BE243845	Hs.75511	connective tissue growth factor
25	102012	102012	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
	102164	102164	NM_000107	Hs.77602	damage-specific DNA binding protein 2 (4
	102283	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	102564	102564	U59423	Hs.79067	MAD (mothers against decapentaplegic, Dr
	102759	102759	NM_005100	Hs.788	A kinase (PRKA) anchor protein (gravin)
30	102804	102804	NM_002318	Hs.83354	lysyl oxidase-like 2
	102898	102898	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,
	103036	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	103095	103095	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and
	103166	103166	AA159248	Hs.180909	peroxiredoxin 1
35	103280	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	103850	103850	AA187101	Hs.213194	hypothetical protein MGC10895
	104592	104592	AW630488	Hs.25338	protease, serine, 23
	104786	104786	AA027167	Hs.10031	KIAA0955 protein
	104865	104865	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
40	104952	104952	AW076098	Hs.345588	desmoplakin (DPI, DPII)
	105178	105178	AA313825	Hs.21941	AD036 protein
	105330	105330	AW338625	Hs.22120	ESTs
	105729	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	105977	105977	AK001972	Hs.30822	hypothetical protein FLJ11110
45	106031	106031	X64116	Hs.171844	Homo sapiens cDNA: FLJ22296 fts, clone H
	106155	106155	AA425414	Hs.33287	nuclear factor I/B
	106423	106423	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15
	107174	107174	BE122762	Hs.25338	ESTs
	107295	107295	AA186629	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
50	108756	108756	AA127221	Hs.117037	ESTs
	108888	108888	AA135606	Hs.189384	gb:z10a05.s1 Soares_pregnant_uterus_NbH
	109166	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabidnes
	109768	109768	F06838	Hs.14763	ESTs
	110906	110906	AA035211	Hs.17404	ESTs
55	111006	111006	BE387014	Hs.166146	Homer, neuronal immediate early gene, 3
	111133	111133	AW580939	Hs.97199	complement component C1q receptor
	113073	113073	N39342	Hs.103042	microtubule-associated protein 1B
	113923	113923	AW953484	Hs.3849	hypothetical protein FLJ22041 similar to
	115061	115061	AI751438	Hs.41271	Homo sapiens mRNA full length insert cDN
60	115145	115145	AA740907	Hs.88297	ESTs
	115947	115947	R47479	Hs.94761	KIAA1691 protein
	116339	116339	AK000290	Hs.44033	dipeptidyl peptidase 8
	116589	116589	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	117023	117023	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
65	117563	117563	AF055634	Hs.44553	unc5 (C.elegans homolog) c
	118475	118475	N66845		gb:za46c1.1.s1 Soares fetal liver spleen
	119073	119073	BE245360	Hs.279477	ESTs
	119174	119174	R71234		gb:ye54c08.s1 Soares placenta Nb2HP Homo
	119416	119416	T97186		gb:ye50h09.s1 Soares fetal liver spleen
70	121335	121335	AA404418		gb:zw37e02.s1 Soares_total_fetus_Nb2HFB_
	123160	123160	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypothe
	123523	123523	AA608588		gb:ae54e06.s1 Stralagene lung carcinoma
	123964	123964	C13961		gb:C13961 Clontech human aorta polyA+ mR
	124315	124315	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene hom
75	124669	124669	AI571594	Hs.102943	hypothetical protein MGC12916
	124875	124875	AI887664	Hs.285814	sprouty (Drosophila) homolog 4
	125103	125103	AA570056	Hs.122730	ESTs, Moderately similar to KIAA1215 pro
	125565	125565	R20840		gb:yg05c08.r1 Soares infant brain 1NIB H

	126511	126511	T92143	Hs.57958	EGF-TM7-tetraphilin-related protein
	126649	126649	AA001860	Hs.279531	ESTs
	449802	449802	AA001860	Hs.279531	ESTs
	127402	127402	AA358869	Hs.227949	SEC13 (S. cerevisiae)-like 1
5	128992	128992	H04150	Hs.107708	ESTs
	129188	129188	NM_001078	Hs.109225	vascular cell adhesion molecule 1
	129371	129371	X06828	Hs.110802	von Willebrand factor
	129765	129765	M86933	Hs.1238	amelogenin (Y chromosome)
	129884	129884	AF055581	Hs.13131	lysosomal
10	130639	130639	AJ557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	130828	130828	AW631469	Hs.203213	ESTs
	131080	131080	NM_001955	Hs.2271	endothelin 1
	131182	131182	AJ824144	Hs.23912	ESTs
	131573	131573	AA040311	Hs.28959	ESTs
15	131756	131756	AA443966	Hs.31595	ESTs
	131881	131881	AW361018	Hs.3383	upstream regulatory element binding prot
	132083	132083	BE386490	Hs.279663	Pirin
	132358	132358	NM_003542	Hs.46423	H4 histone family, member G
	132456	132456	AB011084	Hs.48924	KIAA0512 gene product ALEX2
20	132676	132676	N92589	Hs.261038	ESTs, Weakly similar to I38022 hypothet
	132718	132718	NM_004600	Hs.554	Sjogren syndrome antigen A2 (60kD, ribon
	132760	132760	AA125985	Hs.56145	thymosin, beta, identified in neuroblast
	132968	132968	AF234532	Hs.61638	myosin X
	133081	133081	AJ186431	Hs.296638	prostate differentiation factor
25	133161	133161	AW021103	Hs.6631	hypothetical protein FLJ20373
	133260	133260	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone R
	133491	133491	BE619053	Hs.170001	eukaryotic translation initiation factor
	133550	133550	AJ129903	Hs.74669	vesicle-associated membrane protein 5 (m
	133614	133614	NM_003003	Hs.75232	SEC14 (S. cerevisiae)-like 1
30	133691	133691	M85289	Hs.211573	heparan sulfate proteoglycan 2 (perlecan
	133913	133913	AJ076964	Hs.7753	calumenin
	133985	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec
	134088	134088	AJ379954	Hs.79025	KIAA0096 protein
35	134299	134299	AW580939	Hs.97199	complement component C1q receptor
	116470	116470	AJ272141	Hs.83484	SRY (sex determining region Y)-box 4
	134989	134989	AW968058	Hs.92381	nudix (nucleoside diphosphate linked moi
	135073	135073	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
	100114	100114	X02308	Hs.82962	thymidylate synthetase
40	100143	100143	AJ076465	Hs.278441	KIAA0015 gene product
	100208	100208	NM_002933	Hs.78224	ribonuclease, RNase A family, 1 (pancrea
	100405	100405	AW291587	Hs.82733	nldogen 2
	100455	100455	AW888941	Hs.75789	N-myc downstream regulated
	100618	100618	AJ752163	Hs.114599	collagen, type VIII, alpha 1
45	100658	100658	U56725	Hs.180414	heat shock 70kD protein 2
	100718	100718	BE295928	Hs.75424	inhibitor of DNA binding 1, dominant neg
	100828	100828	AL048753	Hs.303649	small inducible cytokine A2 (monocyte ch
	100991	100991	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	101110	101110	AJ439011	Hs.86386	myeloid cell leukemia sequence 1 (BCL2-r
	101156	101156	AA340987	Hs.75693	prolylcarboxypeptidase (angiotensinase C
50	101184	101184	NM_001674	Hs.460	activating transcription factor 3
	101317	101317	L42176	Hs.8302	four and a half LIM domains 2
	101345	101345	NM_005795	Hs.152175	calcitonin receptor-like
	101475	101475	BE410405	Hs.76288	calpain 2, (mII) large subunit
	101496	101496	X12784	Hs.119129	collagen, type IV, alpha 1
55	101543	101543	M31166	Hs.2050	pentactin-related gene, rapidly induced b
	101560	101560	AW958272	Hs.347326	intercellular adhesion molecule 2
	101592	101592	AF064853	Hs.91299	guanine nucleotide binding protein (G pr
	101634	101634	AV650262	Hs.75765	GRO2 oncogene
	101682	101682	AF043045	Hs.81008	filamin B, beta (actin-binding protein-2
60	101720	101720	M69043	Hs.81328	nuclear factor of kappa light polypeptid
	101744	101744	AJ879352	Hs.118625	hexokinase 1
	101837	101837	M92843	Hs.343586	zinc finger protein homologous to Zfp-36
	101840	101840	AA236291	Hs.183583	serine (or cysteine) proteinase inhibito
65	101864	101864	BE392588	Hs.75777	transgelin
	101966	101966	X98438	Hs.76095	immediate early response 3
	102013	102013	BE616287	Hs.178452	catenin (cadherin-associated protein), a
	102059	102059	AJ752666	Hs.76669	nicotinamide N-methyltransferase
	102283	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	102378	102378	AJ076887	Hs.28491	spermidine/spermine N1-acetyltransferase
70	102460	102460	U48959	Hs.211582	myosin, light polypeptide kinase
	102499	102499	BE243877	Hs.76941	ATPase, Na+K+ transporting, beta 3 poly
	102560	102560	R97457	Hs.63984	cadherin 13, H-cadherin (heart)
	102589	102589	AJ076728	Hs.8867	cysteine-rich, angiogenic inducer, 61
	102645	102645	AL119566	Hs.6721	lysosomal
75	102693	102693	AA532780	Hs.183684	eukaryotic translation initiation factor
	102759	102759	NM_005100	Hs.788	A kinase (PRKA) anchor protein (gravin)

	102882	102882	AI767736	Hs.290070	gelsolin (amyloidosis, Finnish type)
	102915	102915	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin)
	102960	102960	AI904738	Hs.76053	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
5	103020	103020	X53416	Hs.195464	filamin A, alpha (actin-binding protein)
	103036	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial)
	103080	103080	AU077231	Hs.82932	cyclin D1 (PRAD1; parathyroid adenomas)
	103138	103138	X65965		gb.H.sapiens SOD-2 gene for manganese su
	103195	103195	AA351647	Hs.2642	eukaryotic translation elongation factor
	103371	103371	X91247	Hs.13048	thioredoxin reductase 1
10	103471	103471	Y00815	Hs.75216	protein tyrosine phosphatase, receptor t
	104447	104447	AW204145	Hs.156044	ESTs
	104783	104783	AA533513	Hs.93659	protein disulfide isomerase related prot
	104865	104865	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
	104894	104894	AF065214	Hs.18858	phospholipase A2, group IVC (cytosolic,
15	105113	105113	AB037816	Hs.8982	Homo sapiens, clone IMAGE:3506202, mRNA,
	105196	105196	W84893	Hs.9305	angiotensin receptor-like 1
	105263	105263	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	105330	105330	AW338625	Hs.22120	ESTs
	105492	105492	AJ805717	Hs.289112	CGI-43 protein
20	105594	105594	AB024334	Hs.25001	tyrosine 3-monooxygenase/tryptophan 5-mo
	105732	105732	AW504170	Hs.274344	hypothetical protein MGC12942
	105882	105882	W46802	Hs.81988	disabled (Drosophila) homolog 2 (mitogen
	106031	106031	X64116	Hs.171844	Homo sapiens cDNA: FLJ22296 fls, clone H
	106222	106222	AA356392	Hs.21321	Homo sapiens clone FLB9213 PRO2474 mRNA,
25	106263	106263	W21493	Hs.28329	hypothetical protein FLJ14005
	106366	106366	AA186715	Hs.336429	RIKEN cDNA 9130422N19 gene
	106634	106634	W25491	Hs.288909	hypothetical protein FLJ22471
	106793	106793	H94997	Hs.16450	ESTs
	106842	106842	AF124251	Hs.26054	novel SH2-containing protein 3
30	106890	106890	AA489245	Hs.88500	mitogen-activated protein kinase 8 inter
	106974	106974	AJ817130	Hs.9195	Homo sapiens cDNA FLJ13698 fls, clone PL
	107061	107061	BE147611	Hs.6354	stromal cell derived factor receptor 1
	107216	107216	D51069	Hs.211579	melanoma cell adhesion molecule
	107444	107444	W28391	Hs.343258	proliferation-associated 2G4, 38kD
35	108507	108507	AJ554545	Hs.68301	ESTs
	108931	108931	AA147186		gb:z38d01.s1 Stratagene endothelial cel
	109195	109195	AF047033	Hs.132904	solute carrier family 4, sodium bicarbon
	109456	109456	AW956580	Hs.42699	ESTs
	110411	110411	AW001579	Hs.9645	Homo sapiens mRNA for KIAA1741 protein,
40	110906	110906	AA035211	Hs.17404	ESTs
	111091	111091	AA300067	Hs.33032	hypothetical protein DKFZp434N185
	111378	111378	AW160893	Hs.326292	hypothetical gene DKFZp434A1114
	111769	111769	AW629414	Hs.24230	ESTs
	112951	112951	AA307634	Hs.6650	vacuolar protein sorting 45B (yeast homo
45	113195	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	113542	113542	H43374	Hs.7890	Homo sapiens mRNA for KIAA1671 protein,
	113847	113847	NM_005032	Hs.4114	plastin 3 (T isoform)
	113947	113947	W84768		gb:zh53d03.s1 Scarses_fetal_liver_spleen_
	115061	115061	AI751438	Hs.41271	Homo sapiens mRNA full length insert cDN
50	115870	115870	NM_005985	Hs.48029	snail 1 (drosophila homolog), zinc finger
	116228	116228	AI767947	Hs.50841	ESTs
	116314	116314	AI799104	Hs.178705	Homo sapiens cDNA FLJ11333 fls, clone PL
	117023	117023	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
	117156	117156	W73853		ESTs
55	117280	117280	M18217	Hs.172129	Homo sapiens cDNA: FLJ21409 fls, clone C
	119866	119866	AA496205	Hs.193700	Homo sapiens mRNA; cDNA DKFZp586I0324 (f
	121314	121314	W07343	Hs.182538	phospholipid scramblase 4
	121822	121822	AI743860		metallothionein 1E (functional)
60	122331	122331	AL133437	Hs.110771	Homo sapiens cDNA: FLJ21904 fls, clone H
	123160	123160	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti
	124059	124059	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti
	124358	124358	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
	124726	124726	NM_003654	Hs.104576	carbohydrate (keratan sulfate Gal-6) sulf
	125167	125167	AL137540	Hs.102541	netrin 4
65	125307	125307	AW580945	Hs.330466	ESTs
	107985	107985	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
	125598	125598	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
	413731	413731	BE243845	Hs.75511	connective tissue growth factor
	116024	116024	AA088767	Hs.83883	transmembrane, prostate androgen induced
70	418000	418000	AA932794	Hs.83147	guanine nucleotide binding protein-like
	126399	126399	AA088767	Hs.83883	transmembrane, prostate androgen induced
	127566	127566	AI051390	Hs.116731	ESTs
	128453	128453	X02761	Hs.287820	fibronectin 1
75	128515	128515	BE395085	Hs.10086	type I transmembrane protein Fn14
	128623	128623	BE076608	Hs.105509	CTL2 gene
	128669	128669	W28493	Hs.180414	heat shock 70kD protein 8

	128914	128914	AW867491	Hs.107125	plasmalemma vesicle associated protein
	129188	129188	NM_001078	Hs.109225	vascular cell adhesion molecule 1
	129265	129265	AA530892	Hs.171695	dual specificity phosphatase 1
5	129468	129468	AW410538	Hs.111779	secreted protein, acidic, cysteine-rich
	101838	101838	BE243845	Hs.75511	connective tissue growth factor
	129619	129619	AA209534	Hs.284243	tetraspan NET-6 protein
	129762	129762	AA453894	Hs.12372	tripartite motif protein TRIM2
	130018	130018	AA353093		metallothionein 1L
	130178	130178	U20982	Hs.1516	insulin-like growth factor-binding prote
10	130431	130431	AW505214	Hs.155560	calnexin
	130553	130553	AF062649	Hs.252587	pituitary tumor-transforming 1
	130639	130639	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	130686	130686	BE548267	Hs.337886	Homo sapiens cDNA FLJ10934 fis, clone OV
	130818	130818	AW190920	Hs.19928	hypothetical protein SP329
15	130899	130899	AJ077288	Hs.296323	serum glucocorticoid regulated kinase
	131080	131080	NM_001955	Hs.2271	endothelin 1
	131091	131091	AJ271216	Hs.22880	dipeptidylpeptidase III
	131182	131182	AJ824144	Hs.23912	ESTs
	131319	131319	NM_003155	Hs.25590	stanniocalcin 1
20	131328	131328	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131328	131328	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131555	131555	T47364	Hs.278613	Interferon, alpha-inducible protein 27
	131573	131573	AA040311	Hs.28959	ESTs
	131756	131756	AA443966	Hs.31595	ESTs
25	131909	131909	NM_016558	Hs.274411	SCAN domain-containing 1
	132046	132046	AJ359214	Hs.179260	chromosome 14 open reading frame 4
	132151	132151	BE379499	Hs.173705	Homo sapiens cDNA: FLJ22050 fis, clone H
	132187	132187	AA235709	Hs.4193	DKFZP586O1624 protein
	132314	132314	AF112222	Hs.323806	plinin, desmosome associated protein
30	132398	132398	AA876816	Hs.16979	ESTs, Weakly similar to A43932 mucin 2 p
	132490	132490	NM_001290	Hs.4980	LJM domain binding 2
	132548	132548	M24283	Hs.168383	intercellular adhesion molecule 1 (CD54)
	132716	132716	BE378595	Hs.283738	casein kinase 1, alpha 1
	132883	132883	AA373314	Hs.5897	Homo sapiens mRNA; cDNA DKFZp586P1622 (f
35	132989	132989	AA480074	Hs.331328	hypothetical protein FLJ13213
	133071	133071	BE384932	Hs.64313	ESTs, Weakly similar to AF257182 1 G-pro
	133099	133099	W16518	Hs.279518	amyloid beta (A4) precursor-like protein
	133149	133149	AA370045	Hs.6607	AXIN1 up-regulated
40	133200	133200	AB037715	Hs.183639	hypothetical protein FLJ10210
	133260	133260	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone R
	133349	133349	AW631255	Hs.8110	L-3-hydroxyacyl-Coenzyme A dehydrogenase
	133398	133398	NM_000499	Hs.72912	cytochrome P450, subfamily I (aromatic c
	133454	133454	BE547647	Hs.177781	hypothetical protein MGC5618
45	133491	133491	BE619053	Hs.170001	eukaryotic translation initiation factor
	133517	133517	NM_000165	Hs.74471	gap junction protein, alpha 1, 43kD (con
	133538	133538	NM_003257	Hs.74614	tight junction protein 1 (zona occludens
	133584	133584	D90209	Hs.181243	activating transcription factor 4 (tax-r
	133617	133617	BE244334	Hs.75249	ADP-ribosylation factor-like 6 interacti
	133671	133671	AW503116	Hs.301819	zinc finger protein 146
50	133681	133681	AI352558		tyrosine 3-monooxygenase/tryptophan 5-mo
	133730	133730	BE242779	Hs.179526	upregulated by 1,25-dihydroxyvitamin D-3
	133802	133802	AW239400	Hs.76287	G protein-coupled receptor kinase 6
	133838	133838	BE222494	Hs.180919	inhibitor of DNA binding 2, dominant neg
	133889	133889	U48959	Hs.211582	myosin, light polypeptide kinase
55	133975	133975	C18356	Hs.295944	tissue factor pathway inhibitor 2
	134039	134039	NM_002290	Hs.78672	laminin, alpha 4
	134081	134081	AL034349	Hs.79005	protein tyrosine phosphatase, receptor t
	134203	134203	AA161219	Hs.799	diphtheria toxin receptor (heparin-bindi
	134299	134299	AW580939	Hs.97199	complement component C1q receptor
60	134339	134339	R70429	Hs.81988	disabled (Drosophila) homolog 2 (mitogen
	134381	134381	AI557280	Hs.184270	capping protein (actin filament) muscle
	134416	134416	X68264	Hs.211579	melanoma cell adhesion molecule
	134558	134558	NM_001773	Hs.85289	CD34 antigen
65	134983	134983	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p
	135052	135052	AL136653	Hs.93675	decidual protein induced by progesterone
	135069	135069	AA876372	Hs.93961	Homo sapiens mRNA; cDNA DKFZp667D095 (fr
	135073	135073	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
	135196	135196	C03577	Hs.9615	myosin regulatory light chain 2, smooth
	134404	134404	AB000450	Hs.82771	vacuolar related kinase 2
70	100082	100082	AA130080	Hs.4295	proteasome (prosome, macropain) 26S subu
	130150	130150	BE094848	Hs.15113	homogenisate 1,2-dioxygenase (homogeni
	130839	130839	AB011169	Hs.20141	similar to S. cerevisiae SSM4
	100113	100113	NM_001269	Hs.84746	chromosome condensation 1
	100129	100129	AA469369	Hs.5831	tissue inhibitor of metalloproteinase 1
75	100169	100169	AL037228	Hs.82043	D123 gene product
	100190	100190	M91401	Hs.178658	RAD23 (S. cerevisiae) homolog B

	100211	100211	D26528	Hs.123058	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
	130283	130283	NM_012288	Hs.153954	TRAM-like protein
	100248	100248	NM_015156	Hs.78398	KIAA0071 protein
	100262	100262	D38500	Hs.278468	postmeiotic segregation increased 2-like
5	100281	100281	AF091035	Hs.184627	KIAA0118 protein
	100327	100327	D55640		gbHuman monocyte PABL (pseudautosomal
	134495	134495	D63477	Hs.84087	KIAA0143 protein
	135152	135152	M96954	Hs.182741	TIA1 cytotoxic granule-associated RNA-bi
10	100372	100372	NM_014791	Hs.184339	KIAA0175 gene product
	100394	100394	D84284	Hs.66052	CD38 antigen (p45)
	100418	100418	D86978	Hs.84790	KIAA0225 protein
	134347	134347	AF164142	Hs.82042	solute carrier family 23 (nucleobase tra
	100438	100438	AA013051	Hs.91417	topoisomerase (DNA) II binding protein
	100481	100481	X70377	Hs.121489	cystatin D
15	100591	100591	NM_004091	Hs.231444	Homo sapiens, Similar to hypothetical pr
	100662	100662	A1368680	Hs.816	SRY (sex determining region Y)-box 2
	100905	100905	L12260	Hs.172816	neuregulin 1
	100950	100950	AF128542	Hs.166848	polymerase (DNA directed), epsilon
20	135407	135407	J04029	Hs.99936	keratin 10 (epidermolytic hyperkeratosis
	131877	131877	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)
	134786	134786	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous
	134078	134078	L08895	Hs.78995	MADS box transcription enhancer factor 2
	134849	134849	BE409525	Hs.902	neurofibromin 2 (bilateral acoustic neur
25	101152	101152	A1984625	Hs.9884	spindle pole body protein
	131687	131687	BE297635	Hs.3069	heat shock 70kD protein 9B (mortalin-2)
	421155	421155	H87879	Hs.102267	lysyl oxidase
	133975	133975	C18356	Hs.295944	tissue factor pathway inhibitor 2
	130155	130155	AA101043	Hs.151254	kallikrein 7 (chymotryptic, stratum com
	132813	132813	BE313625	Hs.57435	solute carrier family 11 (proton-coupled
30	101300	101300	BE535511		transmembrane trafficking protein
	130344	130344	AW250122	Hs.154879	DiGeorge syndrome critical region gene D
	101381	101381	AW675039	Hs.1227	aminolevulinatase, delta-, dehydratase,
	133780	133780	AA567660	Hs.76152	decorin
35	101447	101447	M21305		gbHuman alpha satellite and satellite 3
	101470	101470	NM_000546	Hs.1846	tumor protein p53 (Li-Fraumeni syndrome)
	101478	101478	NM_002890	Hs.758	RAS p21 protein activator (GTPase activa
	133519	133519	AW583062	Hs.74502	chymotrypsinogen B1
	134116	134116	R84694	Hs.78194	cAMP responsive element binding protein
40	130174	130174	M29551	Hs.151531	protein phosphatase 3 (formerly 2B), cat
	132983	132983	M30269		nidogen (enactin)
	101543	101543	M31166	Hs.2050	pentaxin-related gene, rapidly induced b
	101620	101620	S55271	Hs.247930	Epsilon, IgE
	133595	133595	AA393273	Hs.75133	transcription factor 6-like 1 (mitochond
45	101700	101700	D90337	Hs.247916	neutrophilic peptide precursor C
	134246	134246	D28459	Hs.80612	ubiquitin-conjugating enzyme E2A (RAD6 h
	133948	133948	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
	133948	133948	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
	133948	133948	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
50	101812	101812	BE439894	Hs.78991	DNA segment, numerous copies, expressed
	133396	133396	M86326	Hs.72885	azurocidin 1 (cationic antimicrobial pro
	129026	129026	AL120297	Hs.108043	Friend leukemia virus integration 1
	134831	134831	AA853479	Hs.89890	pyruvate carboxylase
	134395	134395	AA456539	Hs.8262	lysosomal
55	101977	101977	AF112213	Hs.184062	putative Rab5-interacting protein
	101998	101998	U01212	Hs.248153	olfactory marker protein
	102007	102007	U02556	Hs.75307	t-complex-associated-testis-expressed 1-
	416658	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara
	135389	135389	U05237	Hs.99872	fetal Alzheimer antigen
60	130145	130145	U34820	Hs.151051	mitogen-activated protein kinase 10
	420269	420269	U72937	Hs.96264	alpha thalassemia/mental retardation syn
	102123	102123	NM_001809	Hs.1594	centromere protein A (17kD)
	102133	102133	AU076845	Hs.155596	BCL2/adenovirus E1B 19kD-interacting pro
	102162	102162	AA450274	Hs.1592	CDC16 (cell division cycle 16, S. cerevi
65	427653	427653	AA159001	Hs.180069	nuclear respiratory factor 1
	102200	102200	AA232362	Hs.157205	branched chain aminotransferase 1, cytos
	102214	102214	U23752	Hs.32964	SRY (sex determining region Y)-box 11
	131319	131319	NM_003155	Hs.25590	stannocalcin 1
	132316	132316	U28831	Hs.44566	KIAA1641 protein
70	134365	134365	AA568906	Hs.82240	syntaxin 3A
	102298	102298	AA382169	Hs.54483	N-myc (and STAT) Interactor
	302344	302344	BE303044	Hs.192023	eukaryotic translation initiation factor
	102367	102367	U39656	Hs.118825	mitogen-activated protein kinase kinase
	102394	102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma
75	128521	128521	AF071076	Hs.112255	nucleoporin 98kD
	102251	102251	NM_004398	Hs.41706	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
	133746	133746	AW410035	Hs.75862	MAD (mothers against decapentaplegic, Dr

	132828	132828	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-ind
	132828	132828	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-ind
	130441	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic
	129350	129350	U50535	Hs.110630	Human BRCA2 region, mRNA sequence CG006
5	130457	130457	AB014595	Hs.155978	cullin 4B
	102560	102560	R97457	Hs.63984	cadherin 13, H-cadherin (heart)
	134305	134305	U61397	Hs.81424	ubiquitin-like 1 (sentrin)
	132736	132736	AW081883	Hs.211578	Homo sapiens cDNA: FLJ23037 fis, clone L
10	102663	102663	NM_002270	Hs.168075	karyopherin (importin) beta 2
	102735	102735	AF111106	Hs.3382	protein phosphatase 4, regulatory subunit
	101175	101175	U82671	Hs.36980	melanoma antigen, family A, 2
	132164	132164	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio
	102826	102826	NM_007274	Hs.8679	cytosolic acyl coenzyme A thioester hydr
15	102846	102846	BE264974	Hs.6566	thyroid hormone receptor interactor 13
	134161	134161	AA634543	Hs.79440	IGF-II mRNA-binding protein 3
	302363	302363	AW163799	Hs.198365	2,3-bisphosphoglycerate mutase
	125701	125701	T72104	Hs.83194	apolipoprotein A-I
	134656	134656	AI750878	Hs.87409	thrombospondin 1
20	102968	102968	AIU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase
	134037	134037	AI808780	Hs.227730	Integrin, alpha 6
	103023	103023	AW500470	Hs.117950	multifunctional polypeptide similar to S
	130282	130282	BE245380	Hs.153952	5' nucleotidase (CD73)
	128568	128568	H12912	Hs.274691	adenylate kinase 3
25	103093	103093	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine
	129063	129063	X63094	Hs.283822	Rhesus blood group, D antigen
	133227	133227	AW977263	Hs.68257	general transcription factor IIF, polype
	103184	103184	U43143	Hs.74049	fms-related tyrosine kinase 4
	103208	103208	AW411340	Hs.31314	retinoblastoma-binding protein 7
30	131488	131488	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	103334	103334	NM_001260	Hs.25283	cyclin-dependent kinase 8
	135094	135094	NM_003304	Hs.250687	transient receptor potential channel 1
	103352	103352	H09365	Hs.78853	uracil-DNA glycosylase
	132173	132173	X89426	Hs.41716	endothelial cell-specific molecule 1
35	131584	131584	AA598509	Hs.29117	purine-rich element binding protein A
	103378	103378	AL119690	Hs.153618	HCGV11-1 protein
	103410	103410	AA158294	Hs.295362	DR1-associated protein 1 (negative cofac
	103438	103438	AW175781	Hs.152720	M-phase phosphoprotein 6
40	103452	103452	NM_006936	Hs.85119	SMT3 (suppressor of mit two 3, yeast) ho
	135185	135185	AW404908	Hs.96038	Ric (Drosophila)-like, expressed in many
	134662	134662	NM_007048	Hs.284283	butyrophilin, subfamily 3, member A1
	103500	103500	AW408009	Hs.22580	alkylglycerone phosphate synthase
	132084	132084	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)
	133152	133152	Z11695	Hs.324473	mitogen-activated protein kinase 1
45	103612	103612	BE336654	Hs.70937	H3 histone family, member A
	103692	103692	AW137912	Hs.227583	Homo sapiens chromosome X map Xp11.23 L-
	129796	129796	BE218319	Hs.5807	GTPase Rab14
	132683	132683	BE264633	Hs.143638	WD repeat domain 4
	103723	103723	BE274312	Hs.214783	Homo sapiens cDNA FLJ14041 fis, clone HE
50	133260	133260	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone R
	103766	103766	AI920783	Hs.191435	ESTs
	132051	132051	AA393968	Hs.180145	HSPC030 protein
	135289	135289	AW372569	Hs.9788	hypothetical protein MGC10924 similar to
	103794	103794	AF244135	Hs.30670	hepatocellular carcinoma-associated anti
55	134319	134319	BE304999	Hs.285754	fumarate hydratase
	191959	191959	AF142419	Hs.15020	homolog of mouse quaking QKI (KH domain
	103850	103850	AA187101	Hs.213194	hypothetical protein MGC10895
	322026	322026	AW024973	Hs.283675	NPD009 protein
	103861	103861	AA206236	Hs.4944	hypothetical protein FLJ12783
60	447735	447735	AA775268	Hs.6127	Homo sapiens cDNA: FLJ23020 fis, clone L
	131236	131236	AF043117	Hs.24584	ubiquitination factor E4B (homologous to
	129013	129013	AA371156	Hs.107942	DKFZP564M112 protein
	103988	103988	AA314389	Hs.342849	ADP-ribosylation factor-like 5
	425284	425284	AF155568	Hs.348043	NS1-associated protein 1
65	133281	133281	AK001601	Hs.69594	high-mobility group 20A
	108154	108154	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain
	135073	135073	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
	128593	128593	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f
	132084	132084	AA121098	Hs.3838	serum-inducible kinase
70	131427	131427	AF151879	Hs.26706	CGI-121 protein
	104282	104282	C14448	Hs.332338	EST
	130443	130443	D25216	Hs.155650	KIAA0014 gene product
	132837	132837	AA370362	Hs.57958	EGF-TM7-latrophilin-related protein
	104334	104334	D82614	Hs.78771	phosphoglycerate kinase 1
	134731	134731	D89377	Hs.89404	msh (Drosophila) homeo box homolog 2
75	131670	131670	H03514	Hs.15589	ESTs
	104402	104402	H56731	Hs.132956	ESTs

	129077	129077	N74724	Hs.108479	ESTs
	134927	134927	L36531	Hs.91296	Integrin, alpha 8
	134498	134498	AW246273	Hs.84131	threonyl-tRNA synthetase
5	104488	104488	N56191	Hs.106511	protocadherin 17
	129214	129214	AL044335	Hs.109526	zinc finger protein 198
	104530	104530	AK001676	Hs.12457	hypothetical protein FLJ10814
	104544	104544	AI091173	Hs.222362	ESTs, Weakly similar to p40 [H.sapiens]
	104567	104567	AA040620	Hs.5672	hypothetical protein AF140225
10	129575	129575	F08282	Hs.278428	progesterone induced protein
	104599	104599	AW815036	Hs.151251	ESTs
	104667	104667	AI239923	Hs.63931	ESTs
	104764	104764	AI039243	Hs.278585	ESTs
	104787	104787	AA027317		gb:ze97d11.s1 Soares_fetal_heart_NbHH19W
15	104804	104804	AI858702	Hs.31803	ESTs, Weakly similar to N-WASP [H.sapien]
	130828	130828	AW631469	Hs.203213	ESTs
	104943	104943	AF072873	Hs.114218	frizzled (Drosophila) homolog 6
	105024	105024	AA126311	Hs.9879	ESTs
	105038	105038	AW503733	Hs.9414	KIAA1488 protein
20	105096	105096	AL042506	Hs.21589	Kruppel-like factor 7 (ubiquitous)
	105169	105169	BE245294	Hs.180789	S164 protein
	130401	130401	BE396283	Hs.173987	eukaryotic translation initiation factor
	130114	130114	AA233393	Hs.14992	hypothetical protein FLJ11151
	105337	105337	AI468789	Hs.347187	myotubularin related protein 1
25	105376	105376	AW994032	Hs.8768	hypothetical protein FLJ10849
	131962	131962	AK000046	Hs.343877	hypothetical protein FLJ20039
	128658	128658	BE397354	Hs.324830	diphtheria toxin resistance protein requi
	105508	105508	AA173942	Hs.326416	Homo sapiens mRNA; cDNA DKFZp564H1916 (f
	135172	135172	AB028956	Hs.12144	KIAA1033 protein
30	132542	132542	AL137751	Hs.263671	Homo sapiens mRNA; cDNA DKFZp434I0812 (f
	105659	105659	AA283044	Hs.25625	hypothetical protein FLJ11323
	105674	105674	AI609530	Hs.279789	histone deacetylase 3
	105722	105722	AI922821	Hs.32433	ESTs
	115951	115951	BE546245	Hs.301048	sec13-like protein
35	105985	105985	AA406610		gb:zv15b10.s1 Soares_NhHMPu_S1 Homo sapi
	131216	131216	AI815486	Hs.243901	Homo sapiens cDNA FLJ20738 fis, clone HE
	113689	113689	AB037850	Hs.16621	DKFZP434I116 protein
	130839	130839	AB011169	Hs.20141	similar to S. cerevisiae SSM4
	130777	130777	AW135049	Hs.26285	Homo sapiens cDNA FLJ10643 fis, clone NT
40	106196	106196	AA525993	Hs.173699	ESTs, Weakly similar to ALU1_HUMAN ALU S
	133200	133200	AB037715	Hs.183639	hypothetical protein FLJ10210
	106328	106328	AL079559	Hs.28020	KIAA0766 gene product
	106423	106423	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15
	439608	439608	AW864696	Hs.301732	hypothetical protein MGC5306
45	106503	106503	AB033042	Hs.29679	cofactor required for Sp1 transcriptiona
	106543	106543	AA876939	Hs.69285	neuropilin 1
	106589	106589	AK000933	Hs.28661	Homo sapiens cDNA FLJ10071 fis, clone HE
	106596	106596	AA452379		ESTs, Moderately similar to ALU7_HUMAN A
	106636	106636	AW958037	Hs.286	ribosomal protein L4
50	131353	131353	AW754182		gb:RC2-CT0321-131199-011-c01 CT0321 Homo
	131710	131710	NM_015368	Hs.30985	pannexin 1
	131775	131775	AB014548	Hs.31921	KIAA0648 protein
	106773	106773	AA478109	Hs.188833	ESTs
	106817	106817	D61216	Hs.18672	ESTs
55	106848	106848	AA449014	Hs.121025	chromosome 11 open reading frame 5
	418699	418699	BE539639	Hs.173030	ESTs, Weakly similar to ALU8_HUMAN ALU S
	130638	130638	AW021276	Hs.17121	ESTs
	107059	107059	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E. coli Re
	107115	107115	BE379623	Hs.27693	peptidylprolyl isomerase (cyclophilin)-I
60	107156	107156	AA137043	Hs.9663	programmed cell death 6-interacting prot
	130621	130621	AW513087	Hs.16803	LUC7 (S. cerevisiae)-like
	132626	132626	AW504732	Hs.21275	hypothetical protein FLJ11011
	131610	131610	AA357879	Hs.29423	scavenger receptor with C-type lectin
	107295	107295	AA186629	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
65	107315	107315	AA316241	Hs.90691	nucleophosmin/nucleoplasmn 3
	107328	107328	AW959891	Hs.76591	KIAA0887 protein
	134715	134715	U48263	Hs.89040	prepronocleptin
	129938	129938	AW003668	Hs.135587	Human clone 23629 mRNA sequence
70	130074	130074	AL038596	Hs.250745	polymerase (RNA) III (DNA directed) (62k
	132036	132036	AL157433	Hs.37706	hypothetical protein DKFZp434E2220
	113857	113857	AW243158	Hs.5297	DKFZP564A2416 protein
	130419	130419	AF037448	Hs.155489	NS1-associated protein 1
	132616	132616	BE262677	Hs.283558	hypothetical protein PRO1855
	132358	132358	NM_003542	Hs.46423	H4 histone family, member G
75	125827	125827	NM_003403	Hs.97496	YY1 transcription factor
	107609	107609	R75654	Hs.164797	hypothetical protein FLJ13693
	107714	107714	AA015761	Hs.60642	ESTs

	107832	107832	AA021473		gbze66c11.s1 Soares retina N2b4HR Homo
	124337	124337	N23541	Hs.281561	Homo sapiens cDNA: FLJ23582 fs, clone L
	129577	129577	N75346	Hs.308121	CDC20 (cell division cycle 20, S. cerevi
	132000	132000	AW247017	Hs.36978	melanoma antigen, family A, 3
5	107935	107935	AA029428	Hs.61555	ESTs
	131461	131461	AA992841	Hs.27263	KIAA1458 protein
	108029	108029	AA040740	Hs.62007	ESTs
	108084	108084	AA058944	Hs.116602	Homo sapiens, clone IMAGE:4154008, mRNA,
	108168	108168	AA453137	Hs.63176	ESTs
10	108189	108189	AW376061	Hs.63335	ESTs, Moderately similar to A46010 X-tin
	108203	108203	AW847814	Hs.289005	Homo sapiens cDNA: FLJ21532 fs, clone C
	108217	108217	AA058688	Hs.62588	ESTs
	108277	108277	AA064859		gb:zm50103.s1 Stratagene fibroblast (937
	108309	108309	AA069818		gb:zm57e03.r1 Stratagene neuroepithelium
15	108340	108340	AA089820	Hs.180909	peroxiredoxin 1
	108427	108427	AA076382		gb:zm91g08.s1 Stratagene ovarian cancer
	108439	108439	AA078986		gb:zm92h01.s1 Stratagene ovarian cancer
	108469	108469	AA079487		gb:zm97108.s1 Stratagene colon HT29 (937
	108501	108501	AA083256		gb:zm08g12.s1 Stratagene hNT neuron (937
20	108562	108562	AA100796		gb:zm26d06.s1 Stratagene pancreas (93720
	130890	130890	AI907537	Hs.76698	stress-associated endoplasmic reticulum
	130385	130385	AW067800	Hs.155223	stannocalcin 2
	108807	108807	AI652236	Hs.49376	hypothetical protein FLJ20644
	108833	108833	AF188527	Hs.61661	ESTs, Weakly similar to AF174605 1 F-box
25	108846	108846	AL117452	Hs.44155	DKFZP586G1517 protein
	131474	131474	L46353	Hs.2726	high-mobility group (nonhistone chromoso
	108941	108941	AA148650		gb:zo09e06.s1 Stratagene neuroepithelium
	108996	108996	AW995610	Hs.332436	EST
	131183	131183	AI611807	Hs.285107	hypothetical protein FLJ13397
30	109022	109022	AA157291	Hs.21479	ubiquitin 1
	109068	109068	AA164293	Hs.72545	ESTs
	129021	129021	AL044675	Hs.173081	KIAA0530 protein
	109146	109146	AA176589	Hs.142078	EST
	131080	131080	NM_001955	Hs.2271	endothelin 1
35	109222	109222	AA192833	Hs.333512	similar to rat myomegalin
	109481	109481	AA878923	Hs.289069	hypothetical protein FLJ21016
	109516	109516	AI471639	Hs.71913	ESTs
	109556	109556	AI925294	Hs.87385	ESTs
	109578	109578	F02208	Hs.27214	ESTs
40	109625	109625	H29490	Hs.22897	ESTs
	109648	109648	H17800	Hs.7154	ESTs
	109699	109699	H18013	Hs.167483	ESTs
	109933	109933	R52417	Hs.20945	Homo sapiens clone 24993 mRNA sequence
	110039	110039	H11938	Hs.21907	histone acetyltransferase

TABLE 2A

Table 2A shows the accession numbers for those pkays lacking unigenelD's for Table 2. The pkays in Table 7 lacking unigenelD's are represented within Tables 1-6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkay:	Unique Eos probeset Identifier number
CAT number:	Gene cluster number
Accession:	Genbank accession numbers

Pkay	CAT Number	Accession
108469	116761_1	AA079487 AA128547 AA128291 AA079587 AA079600
108501	13684_-12	AA083258
108562	36375_1	AA100796 AF020589 AA074629 AA075948 AA100849 AA085347 AA126309 AA079311 AA079323 AA085274
101300	4669_1	BE535511 M62098 AA306787 AW891768 AA348986 AA338869 AA344013 AW956561 AW389343 AW403607 L40391 AW408435 AA121738 AI568978 H13317 R20373 AW948724 AW948744 AA335023 AA436722 AA448690 C21404 AW884390 AA345454 AA303292 AA174174 BE092290 T90614 AA035104 R76028 AA126924 AA741086 AW022056 AW118940 AA121666 AI832409 AA683475 AI140901 AI623576 AW519064 AW474125 AI953923 AI735349 AW150109 AI436154 AW118130 AW270782 AI804073 NZ7434 AA876543 AA937815 AI051166 AA505378 AI041975 AI335355 AI089540 AA662243 AI127912 AI925604 AI250880 AI366874 AI564386 AI815196 AI683526 AI435885 AI160934 H79030 AI801493 AA448691 AI673767 AI076042 AI804327 AA813438 AA680002 AI274492 T16177 AI287337 AI935050 AA907805 AA911493 AI589411 AI371358 AW576236 AI078866 AW516168 AA346372 AI560185 AA471009 R75857 AA296025 AA523155 AA853168 AI696593 AI658482 AI566601 AW072797 AA128047 AA035502 AW243274 AA992517 R43760
117156	145392_1	W73853 AA828112 W77887 AW889237 AA148524 AI749182 AI754442 AI338392 AI253102 AI079403 AI370541 AI697341 H97538 AW188021 AI927669 W72716 AI051402 AI188071 AI335900 N21488 AW770478 W92522 AI691028 AI913512 AI144448 W73819 AA604358 N28900 W95221 AI868132 H98465 AA148793
125565	1704098_1	R20840 R20839
132983	11922_1	M30269 NM_002508 X82245 AI078760 AW957003 D78945 M27445 AA650439 AL048816 AV660256 AV660347 AA333052 BE295257 T60999 AA383049 AW369677 Z26985 AW175704 AA343326 AW747957 AI818389 W17308 W17302 H15591 AA371284 AA370412 W94966 BE384365 T28498 R80714 R16959 H21723 AW835154 D56097 D56381 W21232 AA190565 AW379755 AW087895
133681	13893_1	AI352558 Z82248 X78138 NM_003405 AU077248 AA223125 S80794 D78577 AI124697 AW403970 BE614089 BE296713 BE621334 L20422 X80536 D54224 D54950 X57345 N29226 AA127798 AA340253 F08031 AA192540 H67636 AA321827 AW950283 AA084159 BE538808 AW401377 AA256774 C03366 W46595 W47608 AA305009 H69431 H69456 AI120082 H11708 AA303717 AA361357 H22042 H78020 AW99584 AA134368 AA322911 AA322961 H60980 N85248 N31547 H79624 T11718 W85828 AW894663 AW894624 BE167441 BE170015 AA304626 AW602163 AW998929 AI156681 AA151067 BE002724 AA608688 H82692 BE155392 AW383636 BE155394 AA487004 AW383504 AI342365 R82553 W16498 BE155344 AI143938 R69901 AA322873 AW340648 R25364 AA367935 AI559406 AA033522 AA374252 AW835019 AI922133 AI697089 N99662 AW189078 AI199076 AW151598 W59944 AA662875 W94022 AA299055 AI039008 AI829449 AA583503 AI635874 AW131665 AI473820 AW273118 AW900930 AA908944 AI688035 AW170272 AI082545 AW468176 AI608761 AI082748 AI911682 AI248943 AI831016 AA192465 AI218477 AA938406 AA385288 AI809817 AA905196 AI191245 AI470204 AI188296 AI421367 AI125315 AI087141 AA629032 AA740589 AI554181 AA150830 AI248541 AI077943 AA775958 AA864930 AI261476 AI123121 AI310394 AA862331 AA872478 BE537084 AI205605 AA720684 AI872093 AW150042 AL120538 AA219627 AA988608 C21397 AI359337 H25337 AI089749 AA605146 AI359620 AA150478 AI359738 AW383642 AW995424 AI766457 R56892 AI089839 W61343 N69107 W46459 AA565955 N20527 AI279782 W46596 AA776573 H23204 AI866231 AI083995 N21530 AA126874 D82630 W65437 AI086917 AW382095 AI086877 H69844 AW340217 W85827 L08439 AA262704 AA505380 W47413 W94135 AA223241 AW089153 AA084101 BE538000 AA096126 T28031 AA491574 R84813 AA774536 AW383522 AA155615 AW383529 AA491520 AW028427 AA171496 AI469689 AW664539 AI811102 AI811116 BE484590 BE350791 H78021 T15405 H21979 AA219489 H13301 AA505883 AI864305 AA423963 AW084401 F04963 R69858 H67097 AI917740 AI655561 H69864 AA033631 AW383484 AI866261 H25293 AA513281 AW271187 H11617 N79982 AI174338 AI904207 AI904208 BE614558 W94127 W65436 AI272249 AA700018 AI579932 AI085941 AW152629
121335	279548_1	AA404418 AI217248
130018	18986_1	AA353093 AW957317 AW872498 AI560785 AI289110 AW135512 X97261 T68873
121822	244391_1	AI743860 N49543 AW027759 BE349467 AI656284 BE463975 R35022 AA370031 AW955302 AL042109 N53092 AI611424 AL079362 AI98290 AI928016 BE394912 BE504220 BE467505 AI611611 AI611407 AI611452 W56437 AI284566 AI583349 AW183058 AI308085 AI074952 AA437315 AA628161 AW301728 AI150224 AA400137 AA437279 AI223355 AA639462 AI261373 AI432414 AI984994 AI539335 AA401550 AA358757 AI609976 AA442357 AA358393 AA437046 AA370301 AA429328 AW272055 AI580502 AI832944 AI038530 AA425107 AI014986 AI148349 AW237721 AW779756 AW137877 AI125293 AA400404 R28554
108309	111495_1	AA069818 AA069971 AA069923 AA069908
107832	genbank_AA021473	AA021473
123523	genbank_AA060858	AA060858
123964	genbank_C13961	C13961
118475	genbank_N66845	N66845
104787	genbank_AA027317	AA027317
106596	304084_1	AI583948 AA578212 AW303715 AA653450 AA456981 AI400385 W88533 AI224133 AW272145 AA088686 R94698
113947	genbank_W84768	W84768
108277	genbank_AA064859	AA064859

WO 02/079492

PCT/US02/04915

	108427	genbank_AA076382	AA076382
	108439	genbank_AA078986	AA078986
	131353	231290_1	AW411259 H23555 AW015049 AJ684275 AW015886 AW068953 AW014085 AJ027260 R52686 AA918278 AJ129462 AA969360 N34869 AJ948416 AA534205 AA702483 AA705292
5	101447	entrez_M21305	M21305
	108931	genbank_AA147186	AA147186
	108941	genbank_AA148650	AA148650
	103138	entrez_X65965	X65965
	119174	genbank_R71234	R71234
10	119416	genbank_T97186	T97186
	105985	genbank_AA406610	AA406610
	100327	entrez_D55640	D55640

TABLE 3:

5	Pkey:	Unique Eos probeset identifier number			
	Accession:	Accession number used for previous patent filings			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UniGeneID:	UniGene number			
	UniGene Title:	UniGene gene title			
10	Pkey	Accession	ExAccn	UniGene	UniGeneTitle
15	100405	D86425	AW291587	Hs.82733	nldogen 2
	100420	D86983	D86983	Hs.118893	Melanoma associated gene
	100481	HG1098-HT1098	X70377	Hs.121489	cystatin D
	100484	HG1103-HT1103	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene hom
	100718	HG3342-HT3519	BE295928	Hs.75424	inhibitor of DNA binding 1, dominant neg
20	100991	J03764	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	101097	L06797	BE245301	Hs.89414	chemokine (C-X-C motif), receptor 4 (fus
	101168	L15388	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	101194	L20971	L20971	Hs.188	phosphodiesterase 4B, cAMP-specific (dun
	101261	L35545	D30857	Hs.82353	protein C receptor, endothelial (EPCR)
25	101345	L76380	NM_005795	Hs.152175	calcitonin receptor-like
	101447	M21305	M21305		gbl/Human alpha satellite and satellite 3
	101485	M24736	AA296520	Hs.89546	selectin E (endothelial adhesion molecu
	101543	M31166	M31166	Hs.2050	pentaxin-related gene, rapidly induced b
	101550	M31551	Y00630	Hs.75716	serine (or cysteine) proteinase inhibito
30	101580	M32334	AW958272	Hs.347326	intercellular adhesion molecule 2
	101674	M61916	NM_002291	Hs.82124	laminin, beta 1
	101714	M68874	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
	101741	M74719	NM_003199	Hs.326198	transcription factor 4
	101838	M92934	BE243845	Hs.75511	connective tissue growth factor
35	101857	M94856	BE550723	Hs.153179	fatty acid binding protein 5 (psoriasis-s
	102012	U03057	BE259035	Hs.118400	linged (Drosophila)-like (sea urchin fas
	102024	U03877	AA301867	Hs.76224	EGF-containing fibulin-like extracellular
	102164	U18300	NM_000107	Hs.77602	damage-specific DNA binding protein 2 (4
	102241	U27109	NM_007351	Hs.268107	multimerin
40	102283	U31384	AW161552	Hs.83381	guanine nucleotide binding protein 11
	102303	U33053	U33053	Hs.2499	protein kinase C-like 1
	102564	U59423	U59423	Hs.79067	MAD (mothers against decapentaplegic, Dr
	102663	U70322	NM_002270	Hs.168075	karyopherin (importin) beta 2
	102759	U81607	NM_005100	Hs.788	A kinase (PRKA) anchor protein (gravin)
45	102778	U83463	AF000652	Hs.8180	syndecan binding protein (syntanin)
	102804	U89942	NM_002318	Hs.83354	lysyl oxidase-like 2
	102887	X04729	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	102898	X06256	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,
	102915	X07820	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
50	103036	X54925	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	103037	X54936	BE018302	Hs.2894	placental growth factor, vascular endoth
	103095	X60957	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and
	103158	X67235	BE242587	Hs.118651	hematopoietically expressed homeobox
	103166	X67951	AA159248	Hs.180909	peroxiredoxin 1
55	103185	X69910	NM_006825	Hs.74368	transmembrane protein (63kD), endoplasm
	103280	X79981	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	103554	Z18951	AI878826	Hs.74034	caveolin 1, caveolae protein, 22kD
	103850	AA187101	AA187101	Hs.213194	hypothetical protein MGC10895
	104465	N24990	Z44203	Hs.26418	ESTs
60	104592	R81003	AW630488	Hs.25338	protease, serine, 23
	104764	AA025351	AI039243	Hs.278585	ESTs
	104788	AA027168	AA027167	Hs.10031	KIAA0955 protein
	104850	AA040465	AL133035	Hs.8728	hypothetical protein DKFZp434G171
	104865	AA045136	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
65	104894	AA054087	AF065214	Hs.18858	phospholipase A2, group IVC (cytosolic,
	104952	AA071089	AW076098	Hs.345588	desmoplakin (DPI, DPII)
	104974	AA085918	Y12059	Hs.278675	bromodomain-containing 4
	105178	AA187490	AA313825	Hs.21941	AD036 protein
	105263	AA227926	AW388633	Hs.6682	solute carrier family 7, (cationic amino
70	105330	AA234743	AW338625	Hs.22120	ESTs
	105376	AA236559	AW994032	Hs.8768	hypothetical protein FLJ10849
	105729	AA292694	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	105826	AA398243	AA478756	Hs.194477	E3 ubiquitin ligase SMURF2
	105977	AA406383	AK001972	Hs.30822	hypothetical protein FLJ11110
75	106008	AA411465	AB033888	Hs.8619	SRY (sex determining region Y)-box 18
	106031	AA412284	X64116	Hs.171844	Homo sapiens cDNA: FLJ22286 fs, clone H
	106124	AA423987	H93366	Hs.7567	Homo sapiens cDNA: FLJ21962 fs, clone H

	106155	AA425309	AA425414	Hs.33287	nuclear factor I/B
	106302	AA435896	AA398859	Hs.18397	hypothetical protein FLJ23221
	106423	AA448238	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15
	106793	AA478778	H94997	Hs.16450	ESTs
5	107174	AA621714	BE122762	Hs.25338	ESTs
	107218	D51069	D51069	Hs.211579	melanoma cell adhesion molecule
	107295	T34527	AA186629	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
	107385	U97519	NM_005397	Hs.16426	podocalyxin-like
10	108756	AA127221	AA127221	Hs.117037	ESTs
	108846	AA132983	AL117452	Hs.44155	DKFZP586G1517 protein
	108888	AA135606	AA135606	Hs.189384	gb:z110a05.s1 Soares_pregnant_uterus_NbH
	109001	AA156125	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to
	109165	AA179845	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines
15	109456	AA232645	AW956580	Hs.42699	ESTs
	109768	F10399	F06838	Hs.14763	ESTs
	110107	H16772	AW151660	Hs.31444	ESTs
	110608	N39584	AA035211	Hs.17404	ESTs
	110984	N52006	AW613287	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
20	111008	N53375	BE387014	Hs.166146	Homer, neuronal immediate early gene, 3
	111018	N54067	AI287912	Hs.3628	mitogen-activated protein kinase kinase
	111133	N64436	AW580939	Hs.97199	complement component C1q receptor
	111760	R26892	BE551929	Hs.268754	Homo sapiens cDNA FLJ11949 fis, clone HE
	113073	T33637	N39342	Hs.103042	microtubule-associated protein 18
25	113195	T57112	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	113923	W80763	AW953484	Hs.3849	hypothetical protein FLJ22041 similar to
	114521	AA046808	AW139036	Hs.108957	40S ribosomal protein S27 isoform
	115061	AA253217	AI751438	Hs.41271	Homo sapiens mRNA full length insert cDN
	115096	AA255991	AI683069	Hs.175319	ESTs
30	115145	AA258138	AA740907	Hs.88297	ESTs
	115819	AA426573	AA486620	Hs.41135	endomucin-2
	115947	AA443793	R47479	Hs.94761	KIAA1691 protein
	116314	AA490588	AI799104	Hs.178705	Homo sapiens cDNA FLJ11333 fis, clone PL
	116339	AA496257	AK000290	Hs.44033	dipeptidyl peptidase 8
35	116430	AA609717	AK001531	Hs.66048	hypothetical protein FLJ10669
	116589	D59570	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	116733	F13787	AL157424	Hs.61289	synaptotagmin 2
	117023	H88157	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp566N0121 (f
	117186	H98988	H98988	Hs.42612	ESTs, Weakly similar to ALU1_HUMAN ALU S
40	117563	N34287	AF055634	Hs.44553	unc5 (C.elegans homolog) c
	117997	N52090	N52090	Hs.47420	EST
	118475	N66845	N66845		gb:z46c11.s1 Soares fetal liver spleen
	118581	N68905	N68905		gb:z46b09.s1 Soares_fetal_lung_NbHL19W
	119073	R32894	BE245360	Hs.279477	ESTs
45	119155	R61715	R61715	Hs.310598	ESTs, Moderately similar to ALU1_HUMAN A
	119174	R71234	R71234		gb:z46c08.s1 Soares placenta Nb2HP Homo
	119221	R98105	C14322	Hs.250700	trypsin beta 1
	119416	T97186	T97186		gb:z46b09.s1 Soares fetal liver spleen
	119866	W80814	AA496205	Hs.193700	Homo sapiens mRNA; cDNA DKFZp566I0324 (f
50	121335	AA404418	AA404418		gb:z46c02.s1 Soares_total_fetus_Nb2HF8_
	121381	AA405747	AW088642	Hs.97984	hypothetical protein FLJ22252 similar to
	123160	AA488687	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti
	123473	AA599143	AA599143		gb:ae52d04.s1 Stratagene lung carcinoma
	123523	AA608588	AA608588		gb:ae54e06.s1 Stratagene lung carcinoma
55	123533	AA608751	AA608751		gb:ae56i07.s1 Stratagene lung carcinoma
	123964	C13961	C13961		gb:C13961 Clontech human aorta polyA+ mR
	124006	D60302	AI147155	Hs.270016	ESTs
	124315	H94892	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene hom
	124659	N93521	AI680737	Hs.289068	Homo sapiens cDNA FLJ11918 fis, clone HE
60	124669	N95477	AI571594	Hs.102943	hypothetical protein MGC12916
	124847	R60044	W07701	Hs.304177	Homo sapiens clone FLB8503 PRO2286 mRNA,
	124875	R70506	AI887664	Hs.285814	sprouty (Drosophila) homolog 4
	125091	T91518	T91518		gb:z46b05.s1 Stratagene lung (937210) H
	125103	T95333	AA570056	Hs.122730	ESTs, Moderately similar to KIAA1215 pro
65	125355	R45630	R60547	Hs.170098	KIAA0372 gene product
	125565	R20839	R20840		gb:yg05c08.r1 Soares infant brain 1N1B H
	125590	R23858	R23858	Hs.143375	Homo sapiens, clone IMAGE:3840937, mRNA,
	125511	AI024874	T92143	Hs.57958	EGF-TM7-latrophilin-related protein
	125563	W26247	AA516391	Hs.181368	US snRNP-specific protein (220 kD), orth
70	126649	AA856990	AA001860	Hs.279531	ESTs
	126872	AA136653	AW450979		gb:U1-H-B13-ala-a-12-0-U1.s1 NCL_CGAP_Su
	127402	AA358869	AA358869	Hs.227949	SEC13 (S. cerevisiae)-like 1
	127651	AI123976	AA362523	Hs.105689	MSTP031 protein
	127759	AI369384	AI369384	Hs.292441	ESTs
75	128062	AA379500	AA379621	Hs.105547	neural proliferation, differentiation an
	128992	R49693	H04150	Hs.107708	ESTs
	129046	AA195678	AB029290	Hs.108258	actin binding protein; macrophilin (microf

	129188	M30257	NM_001076	Hs.109225	vascular cell adhesion molecule 1
	129314	AA028131	BE522768	Hs.290356	mesoderm development candidate 1
	129371	M10321	X06828	Hs.110802	von Willebrand factor
5	129488	J03040	AW410538	Hs.111779	secreted protein, acidic, cysteine-rich
	129765	M86933	M86933	Hs.1238	amelogenin (Y chromosome)
	129805	AA012933	AA012848	Hs.12570	tubulin-specific chaperone d
	129884	AA286710	AF055581	Hs.13131	lysosomal
	130495	AA243278	AW250380	Hs.109059	mitochondrial ribosomal protein L12
10	130639	D59711	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	130657	T94452	AW337575	Hs.201591	ESTs
	130828	AA053400	AW631469	Hs.203213	ESTs
	130972	AA370302	D81866	Hs.21739	Homo sapiens mRNA; cDNA DKFZp586i1518 (f
	131080	J05008	NM_001955	Hs.2271	endothelin 1
15	131137	U85193	W27392	Hs.33287	nuclear factor I/B
	131182	AA256153	AI824144	Hs.23912	ESTs
	131486	X83107	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	131573	AA046593	AA040311	Hs.28959	ESTs
	131647	AA410480	AA359615	Hs.30089	ESTs
20	131756	D45304	AA443966	Hs.31595	ESTs
	131859	M90657	AW960564		transmembrane 4 superfamily member 1
	131881	AA010163	AW361018	Hs.3383	upstream regulatory element binding prot
	132050	AA136353	AI267615	Hs.38022	ESTs
	132083	Y07867	BE386490	Hs.279663	Pirin
25	132164	U84573	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio
	132358	X60486	NM_003542	Hs.46423	H4 histone family, member G
	132413	AA132969	AW361383	Hs.260116	metalloprotease 1 (pittitysin family)
	132456	AA114250	AB011084	Hs.48924	KIAA0512 gene product; ALEX2
	132490	F13782	NM_001290	Hs.4980	LIM domain binding 2
30	132676	AA283035	N92589	Hs.261038	ESTs, Weakly similar to I38022 hypotheti
	132687	AB002301	AB002301	Hs.54985	KIAA0303 protein
	132718	AA056731	NM_004600	Hs.554	Sjogren syndrome antigen A2 (60kD, ribon
	132736	U68019	AW081883	Hs.211578	Homo sapiens cDNA: FLJ23037 fis, clone L
	132760	H99198	AA125985	Hs.56145	thymosin, beta, identified in neuroblast
35	132933	AA598702	BE263252	Hs.6101	hypothetical protein MGC3178
	132968	N77151	AF234532	Hs.61638	myosin X
	132994	AA505133	AA112748	Hs.279905	clone HQ0310 PRO0310p1
	133061	AB000584	AI186431	Hs.296638	prostate differentiation factor
	133147	D12763	AA026533	Hs.66	Interleukin 1 receptor-like 1
40	133161	AA253193	AW021103	Hs.6631	hypothetical protein FLJ20373
	133200	AA432248	AB037715	Hs.183639	hypothetical protein FLJ10210
	133260	AA083572	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone R
	133363	AA479713	AI866286	Hs.71962	ESTs, Weakly similar to B36298 proline-r
	133491	L40395	BE619053	Hs.170001	eukaryotic translation initiation factor
45	133517	X52947	NM_000165	Hs.74471	gap junction protein, alpha 1, 43kD (con
	133550	M80846	AI129903	Hs.74669	vesicle-associated membrane protein 5 (m
	133607	M34539	BE273749		FK506-binding protein 1A (12kD)
	133614	D67029	NM_003003	Hs.75232	SEC14 (S. cerevisiae)-like 1
	133627	U09587	NM_002047	Hs.75280	glycyl-tRNA synthetase
50	133691	M85289	M85289	Hs.211573	heparan sulfate proteoglycan 2 (perlecan
	133696	D10522	AI878921	Hs.75607	myristoylated alanine-rich protein kinas
	133913	W84712	AU076964	Hs.7753	calumenin
	133975	D29992	C18356	Hs.295944	tissue factor pathway inhibitor 2
	133985	L34657	L34657	Hs.78146	platelet/endothelial cell adhesion molec
55	134039	S78569	NM_002280	Hs.78672	laminin, alpha 4
	134088	D43636	AI379954	Hs.79025	KIAA0096 protein
	134161	U97188	AA634543	Hs.79440	IGF-II mRNA-binding protein 3
	134299	AA487558	AW580939	Hs.97199	complement component C1q receptor
	134416	M28882	X68264	Hs.211579	melanoma cell adhesion molecule
60	134453	X70683	AI272141	Hs.83484	SRY (sex determining region Y)-box 4
	134656	X14787	AI750878	Hs.87409	thrombospondin 1
	134989	AA236324	AW968058	Hs.92381	nucleoside diphosphate linked moi
	135051	C15324	AI272141	Hs.83484	SRY (sex determining region Y)-box 4
	135073	AA452000	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
65	135349	D83174	AA114212	Hs.9930	serine (or cysteine) proteinase inhibito
	100114	D00596	X02308	Hs.82962	thymidylate synthetase
	100130	D11428	NM_000304	Hs.103724	peripheral myelin protein 22
	100143	D13640	AU076465	Hs.278441	KIAA0015 gene product
	100168	D14874	H73444	Hs.394	adrenomedullin
70	100208	D26129	NM_002933	Hs.78224	ribonuclease, RNase A family, 1 (pancrea
	100224	D28476	AL121516	Hs.136617	thyroid hormone receptor interactor 12
	100405	D86425	AW291587	Hs.82733	nidogen 2
	100420	D86983	D86983	Hs.118893	Melanoma associated gene
	100455	D87953	AW888941	Hs.75789	N-myc downstream regulated
75	100529	HG1862-HT1897	BE313693	Hs.334330	calmodulin 2 (phosphorylase kinase, delt
	100618	HG2614-HT2710	AI752163	Hs.114599	collagen, type VIII, alpha 1
	100619	HG2639-HT2735	N24433	Hs.241567	RNA binding motif, single stranded inter

	100658	HG2855-HT2995	U56725	Hs.180414	heat shock 70kD protein 2
	100676	HG3044-HT3742	X02761	Hs.287820	fibronectin 1
	100718	HG3342-HT3519	BE295928	Hs.75424	inhibitor of DNA binding 1, dominant neg
	100752	HG3543-HT3739	T81309		insulin-like growth factor 2 (somatomedi
5	100828	HG4089-HT4339	AL048753	Hs.303649	small inducible cytokine A2 (monocyte ch
	100850	HG417-HT417	AA836472	Hs.297939	cathepsin B
	100991	J03764	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	101097	L08797	BE245301	Hs.89414	chemokine (C-X-C motif), receptor 4 (fus
10	101110	L08246	AI439011	Hs.86386	myeloid cell leukemia sequence 1 (BCL2-r
	101142	L12711	L12711	Hs.89643	transketolase (Wernicke-Korsakoff syndro
	101156	L13977	AA340987	Hs.75693	prolylcarboxypeptidase (angiotensinase C
	101168	L15388	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	101184	L19871	NM_001674	Hs.460	activating transcription factor 3
15	101192	L20859	BE247295	Hs.78452	solute carrier family 20 (phosphate tran
	101317	L42176	L42176	Hs.8302	four and a half LIM domains 2
	101336	L49169	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h
	101345	L76380	NM_005795	Hs.152175	calcitonin receptor-like
	101400	M15990	M15990	Hs.194148	v-jes-1 Yamaguchi sarcoma viral oncogene
20	101475	M23254	BE410405	Hs.76288	calpain 2, (mII) large subunit
	101485	M24736	AA296520	Hs.89546	selectin E (endothelial adhesion molecu
	101496	M26576	X12784	Hs.119129	collagen, type IV, alpha 1
	101505	M27396	AA307580	Hs.75692	asparagine synthetase
	101543	M31166	M31166	Hs.2050	pentactin-related gene, rapidly induced b
25	101557	M31994	BE293116	Hs.76392	aldehyde dehydrogenase 1 family, member
	101560	M32334	AW958272	Hs.347326	intercellular adhesion molecule 2
	101587	M35878	AI752416	Hs.77326	insulin-like growth factor binding prote
	101592	M36429	AF064853	Hs.91299	guanine nucleotide binding protein (G pr
	101633	M57730	NM_004428	Hs.1624	ephritin-A1
	101634	M57731	AV650262	Hs.75765	GRO2 oncogene
30	101667	M60858	NM_005381		nucleolin
	101682	M62994	AF043045	Hs.81008	filamin B, beta (actin-binding protein-2
	101714	M68874	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
	101720	M69043	M69043	Hs.81328	nuclear factor of kappa light polypeptid
35	101741	M74719	NM_003199	Hs.326198	transcription factor 4
	101744	M75126	AI879352	Hs.118625	hexokinase 1
	101793	M84349	W01076	Hs.278573	CD59 antigen p18-20 (antigen identified
	101837	M82843	M92843	Hs.343586	zinc finger protein homologous to Zip-36
	101838	M92934	BE243845	Hs.75511	connective tissue growth factor
40	101840	M83056	AA236291	Hs.183583	serine (or cysteine) proteinase inhibito
	101857	M94856	BE550723	Hs.153179	fatty acid binding protein 5 (psoriasis-
	101864	M95787	BE392588	Hs.75777	transgelin
	101931	S76965	NM_006823	Hs.75209	protein kinase (cAMP-dependent, catalyti
	101966	S81914	X96438	Hs.76095	immediate early response 3
45	102012	U03057	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
	102013	U03100	BE618287	Hs.178452	catenin (cadherin-associated protein), a
	102024	U03877	AA301857	Hs.76224	EGF-containing fibulin-like extracellular
	102059	U08021	AI752668	Hs.76669	nicotinamide N-methyltransferase
	102121	U14391	NM_004998	Hs.82251	myosin IE
50	102283	U31384	AW161552	Hs.83381	guanine nucleotide binding protein 11
	102300	U32944	AI929721	Hs.5120	dynein, cytoplasmic, light polypeptide
	102378	U40369	AU076887	Hs.28491	spermidine/spermine N1-acetyltransferase
	102395	U41767	AU077005	Hs.92208	a disintegrin and metalloproteinase doma
	102460	U48959	U48959	Hs.211582	myosin, light polypeptide kinase
55	102491	U51010	U51010		gbxHuman nicotinamide N-methyltransferas
	102499	U51478	BE243877	Hs.76941	ATPase, Na ⁺ /K ⁺ transporting, beta 3 poly
	102523	U53445	U53445	Hs.15432	downregulated in ovarian cancer 1
	102560	U59289	R97457	Hs.63984	cadherin 13, H-cadherin (heart)
	102564	U59423	U59423	Hs.79067	MAD (mothers against decapentaplegic, Dr
60	102589	U62015	AU076728	Hs.8867	cysteine-rich, angiogenic inducer, 61
	102600	U63825	AI984144	Hs.66713	hepatitis delta antigen-interacting prot
	102645	U67963	AL119566	Hs.6721	lysosomal
	102687	U73379	NM_007019	Hs.93002	ubiquitin carrier protein E2-C
	102693	U73824	AA532780	Hs.183684	eukaryotic translation initiation factor
65	102709	U77604	AA122237	Hs.81874	microsomal glutathione S-transferase 2
	102759	U81607	NM_005100	Hs.788	A kinase (PRKA) anchor protein (gravin)
	102804	U89942	NM_002318	Hs.83354	lysyl oxidase-like 2
	102882	X04412	AI767736	Hs.290070	gelsolin (amyloidosis, Finnish type)
	102907	X06985	BE409861	Hs.202833	heme oxygenase (decycling) 1
70	102915	X07820	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	102927	X12876	BE512730	Hs.65114	keratin 18
	102960	X15729	AI904738	Hs.76053	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
	103011	X52541	AJ243425	Hs.326035	early growth response 1
	103020	X53416	X53416	Hs.195464	filamin A, alpha (actin-binding protein-
75	103029	X54489	AW800726	Hs.789	GRO1 oncogene (melanoma growth stimulat
	103036	X54925	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	103056	X57206	Y18024	Hs.78877	inositol 1,4,5-trisphosphate 3-kinase B

	103080	X59798	AU077231	Hs.82932	cyclin D1 (PRAD1: parathyroid adenomas
	103085	X60957	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and
	103138	X65965	X65965		gbHLsapiens SOD-2 gene for manganese su
	103176	X69111	AL021154	Hs.76884	inhibitor of DNA binding 3, dominant neg
5	103185	X70940	AA351647	Hs.2842	eukaryotic translation elongation factor
	103347	X87838	AU077309	Hs.171271	catenin (cadherin-associated protein), b
	103371	X91247	X91247	Hs.13046	thioredoxin reductase 1
	103432	X97748	X97748		gbHLsapiens PTX3 gene promotor region.
	103471	Y00815	Y00815	Hs.75216	protein tyrosine phosphatase, receptor l
10	103967	AA303711	AL120051	Hs.144700	ephrin-B1
	104447	LA4538	AW204145	Hs.156044	ESTs
	104784	AA025351	AJ039243	Hs.278585	ESTs
	104783	AA027050	AA533513	Hs.93659	protein disulfide isomerase related prot
15	104798	AA029462	AW952619	Hs.17235	Homo sapiens clone TCCCA00176 mRNA sequ
	104865	AA045136	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
	104877	AA047437	AJ138635	Hs.22968	Homo sapiens clone IMAGE:451939, mRNA se
	104894	AA054087	AF065214	Hs.18858	phospholipase A2, group IVC (cytosolic,
	104952	AA071089	AW076098	Hs.345588	desmoplakin (DPI, DPII)
20	105113	AA156450	AB037816	Hs.8982	Homo sapiens, clone IMAGE:3506202, mRNA,
	105178	AA187490	AA313825	Hs.21941	AD036 protein
	105196	AA195031	W84893	Hs.9305	angiotensin receptor-like 1
	105215	AA205724	AA205759	Hs.10119	hypothetical protein FLJ14957
	105283	AA227926	AW388633	Hs.6682	solute carrier family 7, (cationic amino
25	105271	AA227986	AA807881	Hs.25329	ESTs
	105330	AA234743	AW338625	Hs.22120	ESTs
	105461	AA253216	BE539071	Hs.69388	hypothetical protein FLJ20505
	105492	AA256210	AI805717	Hs.289112	CGI-43 protein
	105493	AA256268	AL047586	Hs.10283	RNA binding motif protein 8B
30	105594	AA278397	AB024334	Hs.25001	tyrosine 3-monooxygenase/tryptophan 5-mo
	105727	AA292379	AL135159	Hs.20340	KIAA1002 protein
	105732	AA292717	AW504170	Hs.274344	hypothetical protein MGC12942
	105767	AA346551	AW370946	Hs.23457	ESTs
	105882	AA400292	W46802	Hs.81988	disabled (Drosophila) homolog 2 (mitogen
35	105936	AA404338	AI878765	Hs.21812	ESTs
	106031	AA412284	X64116	Hs.171844	Homo sapiens cDNA: FLJ22298 fs, clone H
	106124	AA423987	H93366	Hs.7567	Homo sapiens cDNA: FLJ21962 fs, clone H
	106222	AA428594	AA356392	Hs.21321	Homo sapiens clone FLB9213 PRO2474 mRNA,
	106241	AA430108	BE019681	Hs.6019	Homo sapiens cDNA: FLJ21288 fs, clone C
40	106263	AA431462	W21493	Hs.28329	hypothetical protein FLJ14005
	106264	AA431470	AL046859	Hs.3407	protein kinase (cAMP-dependent, catalyti
	106366	AA443756	AA186715	Hs.336429	RIKEN cDNA 913042Z19 gene
	106454	AA449479	NM_014038	Hs.5216	HSPC028 protein
	106634	AA459916	W25491	Hs.288909	hypothetical protein FLJ22471
45	106724	AA465226	N48670	Hs.28631	Homo sapiens cDNA: FLJ22141 fs, clone H
	106793	AA478778	H94997	Hs.16450	ESTs
	106799	AA479037	BE313412	Hs.7961	Homo sapiens clone 25012 mRNA sequence
	106842	AA482597	AF124251	Hs.28054	novel SH2-containing protein 3
50	106868	AA487561	BE185536	Hs.301183	molecule possessing ankryrin repeats indu
	106890	AA489245	AA489245	Hs.88500	mitogen-activated protein kinase 8 inter
	106961	AA504110	AW243614	Hs.18063	Homo sapiens cDNA FLJ10768 fs, clone NT
	106974	AA520989	AI817130	Hs.9195	Homo sapiens cDNA FLJ13698 fs, clone PL
	107030	AA599434	AL117424	Hs.25035	chloride intracellular channel 4
	107061	AA608649	BE147611	Hs.6354	stromal cell derived factor receptor 1
55	107086	AA609519	NM_012331	Hs.26458	methionine sulfoxide reductase A
	107216	D51069	D51069	Hs.211579	melanoma cell adhesion molecule
	107385	U97519	NM_005397	Hs.16426	podocalyxin-like
	107444	W28391	W28391	Hs.343258	proliferation-associated 2G4, 38kD
	107885	AA035638	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
60	108507	AA083514	AI554545	Hs.68301	ESTs
	108695	AA121315	AB029000	Hs.70823	KIAA1077 protein
	108931	AA147186	AA147186		gb:zo38d01.s1 Stratagene endothelial cel
	109001	AA156125	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to
	109195	AA188932	AF047033	Hs.132904	solute carrier family 4, sodium bicarbon
65	109390	AA219653	AW007485	Hs.87125	EH-domain containing 3
	109456	AA232645	AW956580	Hs.42699	ESTs
	109737	F10078	AA055415	Hs.13233	ESTs, Moderately similar to A47582 B-cell
	110411	H48032	AW001579	Hs.9645	Homo sapiens mRNA for KIAA1741 protein,
	110660	H82117	AA782114	Hs.28043	ESTs
70	110906	N39584	AA035211	Hs.17404	ESTs
	111018	N54067	AI287912	Hs.3628	mitogen-activated protein kinase kinase
	111091	N59858	AA300067	Hs.33032	hypothetical protein DKFZp434N185
	111356	N90933	BE301871	Hs.4867	mannosyl (alpha-1,3)-glycoprotein beta-
	111378	N93764	AW160993	Hs.326292	hypothetical gene DKFZp434A1114
75	111741	R26124	AB020653	Hs.24024	KIAA0846 protein
	111769	R27957	AW629414	Hs.24230	ESTs
	112318	R55470	AW063384	Hs.11087	ESTs, Highly similar to T46395 hypotheti

	112951	T18550	AA307634	Hs.6650	vacuolar protein sorting 45B (yeast homo
	113057	T26674	AW194301	Hs.339283	Human DNA sequence from clone RP1-187J11
	113195	T57112	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
5	113490	T88700	BE178110	Hs.173374	Homo sapiens cDNA FLJ10500 fls, clone NT
	113542	T90527	H43374	Hs.7890	Homo sapiens mRNA for KIAA1671 protein,
	113803	W42789	AW880709	Hs.283683	chromosome 8 open reading frame 4
	113847	W60002	NM_005032	Hs.4114	plastin 3 (T isoform)
	113910	W78175	AA113262	Hs.17901	Homo sapiens, clone IMAGE:3937015, mRNA,
10	113947	W84768	W84768		gb:zh53d03.s1 Soares_fetal_liver_spleen_
	114047	W94427	AL035858	Hs.3807	FXD domain-containing ion transport reg
	115061	AA253217	AI751438	Hs.41271	Homo sapiens mRNA full length insert cDN
	115819	AA426573	AA486620	Hs.41135	endomucin-2
	115870	AA432374	NM_005985	Hs.48029	snail 1 (drosophila homolog), zinc finger
15	115984	AA446622	AA987568	Hs.74313	KIAA1265 protein
	116228	AA478771	AI787947	Hs.50841	ESTs
	116284	AA482594	D51174	Hs.272239	lysosomal
	116314	AA490588	AI799104	Hs.178705	Homo sapiens cDNA FLJ11333 fls, clone PL
	116589	D59570	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
20	117023	H88157	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
	117112	H94648	AW969999	Hs.293658	ESTs
	117156	H97538	W73853		ESTs
	117176	H88670	H45100	Hs.49753	uvect autotantigen with coiled coil domai
	117280	N22107	M18217	Hs.172129	Homo sapiens cDNA: FLJ21409 fls, clone C
25	119559	W38197	W38197		Empirically selected from AFFX single pr
	119866	W80814	AA496205	Hs.193700	Homo sapiens mRNA; cDNA DKFZp586I0324 (f
	120655	AA287347	AA305599	Hs.238205	hypothetical protein PRO2013
	121314	AA402799	W07343	Hs.182538	phospholipid scramblase 4
	121335	AA404418	AA404418		gb:zw37e02.s1 Soares_total_fetus_Nb2HF8_
30	121822	AA425107	AI743860		metallothionein 1E (functional)
	121835	AA425435	AB033030	Hs.300670	KIAA1204 protein
	122331	AA442872	AL133437	Hs.110771	Homo sapiens cDNA: FLJ21904 fls, clone H
	122577	AA452860	AA829725	Hs.334437	hypothetical protein MGCA248
	123160	AA488687	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti
35	123486	AA599674	BE019072	Hs.334802	Homo sapiens cDNA FLJ14680 fls, clone NT
	124069	F13673	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti
	124339	H99093	H99093	Hs.343411	DEADH (Asp-Glu-Ala-Asp/His) box polypep
	124358	N22495	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
	124364	N23031	AF265555	Hs.250646	baculoviral IAP repeat-containing 6
40	124726	R15740	NM_003654	Hs.104576	carbohydrate (keratan sulfate Gal-6) sul
	124763	R39810	BE410405	Hs.76288	calpain 2, (mII) large subunit
	125167	W45560	AL137540	Hs.102541	netrin 4
	125304	Z39833	AL359573	Hs.124940	GTP-binding protein
	125307	Z40583	AW580945	Hs.330466	ESTs
45	125329	AA825437	AA825437	Hs.58875	ESTs
	125598	R68613	T40084	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
	125609	AA868063	AA868063	Hs.104576	carbohydrate (keratan sulfate Gal-6) sul
	126245	AA128075	AA088767	Hs.83883	transmembrane, prostate androgen induced
	127435	N66570	X69086	Hs.286161	Homo sapiens cDNA FLJ13613 fls, clone PL
50	127566	AI051390	AI051390	Hs.116731	ESTs
	127619	AA627122	AA627122	Hs.163787	ESTs
	128453	X02761	X02761	Hs.287820	fibronectin 1
	128495	AF010193	NM_005804	Hs.100602	MAD (mothers against decapentaplegic, Dr
	128515	AA149044	BE395085	Hs.10086	type I transmembrane protein Fn14
55	128580	U82108	U82108	Hs.101813	solute carrier family 9 (sodium/hydrogen
	128623	D78676	BE076508	Hs.105509	CTL2 gene
	128642	L35240	Z28913	Hs.102948	enigma (LIM domain protein)
	128669	AA598737	W28493	Hs.180414	heat shock 70kD protein 8
	128903	R69417	AW150717	Hs.345728	STAT induced STAT inhibitor 3
60	128914	AA232837	AW867491	Hs.107125	plasmalemma vesicle associated protein
	129087	N72695	AI348027	Hs.108557	hypothetical protein PP1057
	129188	M30257	NM_001078	Hs.109225	vascular cell adhesion molecule 1
	129226	M96843	BE222494	Hs.180919	inhibitor of DNA binding 2, dominant neg
	129265	X88277	AA530892	Hs.171695	dual specificity phosphatase 1
65	129345	AA292440	R22497	Hs.110571	growth arrest and DNA-damage-inducible,
	129468	J03040	AW410538	Hs.111779	secreted protein, acidic, cysteine-rich
	129488	AA228107	AW966728	Hs.54642	methionine adenosyltransferase II, beta
	129498	AA449789	AA449789	Hs.75511	connective tissue growth factor
	129557	W01357	AL045404	Hs.46366	KIAA0948 protein
70	129619	AA610116	AA209534	Hs.284243	tetraspan NET-8 protein
	129627	AA258308	T40084	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
	129762	AA460273	AA453694	Hs.12372	tripartite motif protein TRIM2
	129884	AA286710	AF055581	Hs.13131	lysosomal
	130018	T68873	AA353093		metallothionein 1L
75	130147	D63476	D63476	Hs.172813	PAK-interacting exchange factor beta
	130178	M62403	U20982	Hs.1516	Insulin-like growth factor-binding prote
	130282	X55740	BE245380	Hs.153952	5' nucleotidase (CD73)

	130431	L10284	AW505214	Hs.155560	calnexin
	130485	AA243278	AW250380	Hs.109059	mitochondrial ribosomal protein L12
	130553	AA430032	AF062649	Hs.252587	pituitary tumor-transforming 1
	130638	H16402	AW021276	Hs.17121	ESTs
5	130639	D59711	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	130657	T94452	AW337575	Hs.201591	ESTs
	130686	AA431571	BE548267	Hs.337986	Homo sapiens cDNA FLJ10934 fs, clone OV
	130776	R79356	AF167706	Hs.19280	cysteine-rich motor neuron 1
	130818	AA280375	AW190920	Hs.19928	hypothetical protein SP329
10	130840	Z49269	BE048821	Hs.20144	small inducible cytokine subfamily A (Cy
	130899	Z41740	AI077288	Hs.296323	serum/glucocorticoid regulated kinase
	131002	AA121543	AL050295	Hs.22039	KIAA0758 protein
	131080	J05008	NM_001955	Hs.2271	endothelin 1
	131084	AA101878	NM_017413	Hs.303084	apelin; peptide ligand for APJ receptor
15	131091	T35341	AJ271216	Hs.22880	dipterydylpeptidase III
	131107	N87590	BE620886	Hs.75354	GCN1 (general control of amino-acid synt
	131182	AA256153	AI824144	Hs.23912	ESTs
	131207	W74533	AF104266	Hs.24212	leukophilin
	131319	U25997	NM_003155	Hs.25590	stanniocalcin 1
20	131328	V01512	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131328	V01512	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131328	V01512	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131328	V01512	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131509	X56681	X56681	Hs.2780	jun D proto-oncogene
25	131555	AA161292	T47364	Hs.278613	interferon, alpha-inducible protein 27
	131564	AA491465	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fs, clone PL
	131573	AA046593	AA040311	Hs.28959	ESTs
	131692	D50914	BE559681	Hs.30736	KIAA0124 protein
	131756	D45304	AA443966	Hs.31595	ESTs
30	131859	M90657	AW960564		transmembrane 4 superfamily member 1
	131909	W69127	NM_016558	Hs.274411	SCAN domain-containing 1
	131915	AA316186	AI161383	Hs.34549	ESTs, Highly similar to S94541 1 clone 4
	132046	AA384503	AI359214	Hs.179260	chromosome 14 open reading frame 4
	132060	AA136353	AI267615	Hs.38022	ESTs
35	132151	AA044755	BE379499	Hs.173705	Homo sapiens cDNA: FLJ22050 fs, clone H
	132164	U84573	AI752235	Hs.41270	procollagen-llysine, 2-oxoglutarate 5-dio
	132187	AA058911	AA235709	Hs.4193	DKFZP586O1624 protein
	132303	AA620962	BE177330	Hs.325093	Homo sapiens cDNA: FLJ21210 fs, clone C
	132314	AA285290	AF112222	Hs.323806	pinin, desmosome associated protein
40	132358	X60486	NM_003542	Hs.46423	H4 histone family, member G
	132398	R31641	AA876616	Hs.16979	ESTs, Weakly similar to A43932 mucin 2 p
	132421	AA489190	AW163483	Hs.48320	double ring-finger protein, Dorrin
	132490	F13782	NM_001290	Hs.4980	LIM domain binding 2
	132520	AA257993	AA257992	Hs.50651	Janus kinase 1 (a protein tyrosine kinas
45	132546	M24283	M24283	Hs.168383	intercellular adhesion molecule 1 (CD54)
	132610	AA443114	AA160511	Hs.5326	amino acid system N transporter 2; porc
	132716	T35289	BE379595	Hs.283738	casein kinase 1, alpha 1
	132840	N23817	BE218319	Hs.5807	GTPase Rab14
	132883	AA047151	AA373314	Hs.5897	Homo sapiens mRNA; cDNA DKFZp586P1622 (f
50	132968	N77151	AF234532	Hs.61638	myosin X
	132989	AA480074	AA480074	Hs.331328	hypothetical protein FLJ13213
	132999	Y00787	Y00787	Hs.624	Interleukin 8
	133071	T99789	BE384932	Hs.64313	ESTs, Weakly similar to AF257182 1 G-pro
	133076	W84341	AW946276	Hs.6441	Homo sapiens mRNA; cDNA DKFZp586J021 (fr
55	133099	L09209	W16518	Hs.279518	amyloid beta (A4) precursor-like protein
	133147	D12763	AA026533	Hs.66	Interleukin 1 receptor-like 1
	133149	T16484	AA370045	Hs.6607	AXIN1 up-regulated
	133161	AA253193	AW021103	Hs.6631	hypothetical protein FLJ20373
	133200	AA432248	AB037715	Hs.183639	hypothetical protein FLJ10210
60	133220	X82200	NM_006074	Hs.318501	Homo sapiens mRNA full length insert cDN
	133260	AA083572	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fs, clone R
	133295	L00352	AI147861	Hs.213289	low density lipoprotein receptor (familia
	133349	N75791	AW631255	Hs.8110	L-3-hydroxyacyl-Coenzyme A dehydrogenase
	133391	X57579	AW103364	Hs.727	Inhibin, beta A (activin A, activin AB a
65	133398	X02612	NM_000499	Hs.72912	cytochrome P450, subfamily I (aromatic c
	133436	H44631	BE294068	Hs.737	immediate early protein
	133454	AA080257	BE547647	Hs.177781	hypothetical protein MGC5618
	133478	X83703	X83703	Hs.31432	cardiac ankyrin repeat protein
	133491	L40395	BE619053	Hs.170001	eukaryotic translation initiation factor
70	133510	AA227913	AW880841	Hs.96908	p53-induced protein
	133517	X52947	NM_000165	Hs.74471	gap junction protein, alpha 1, 43kD (con
	133526	M11313	AU077051	Hs.74561	alpha-2-macroglobulin
	133538	L14837	NM_003257	Hs.74614	tight junction protein 1 (zona occludens
	133562	M60721	M60721	Hs.74870	H2.0 (Drosophila)-like homeo box 1
75	133584	D90209	D90209	Hs.181243	activating transcription factor 4 (tax-r
	133590	T67986	T70956	Hs.75106	clusterin (complement lysis inhibitor, S

5	133617	AA148318	BE244334	Hs.75249	ADP-ribosylation factor-like 6 interacti
	133651	U97105	AJ301740	Hs.173381	dihydropyrimidinase-like 2
	133671	T25747	AW503116	Hs.301819	zinc finger protein 146
	133678	K02574	AW247252		nucleoside phosphorylase
	133681	D78577	AJ352558		tyrosine 3-monooxygenase/tryptophan 5-mo
	133722	X53331	AW969976	Hs.278009	matrix Gla protein
	133730	S73591	BE242779	Hs.179526	upregulated by 1,25-dihydroxyvitamin D-3
	133750	X95735	BE410769	Hs.75873	zyxin
10	133802	L16862	AW239400	Hs.76297	G protein-coupled receptor kinase 6
	133825	U44975	BE616902	Hs.285313	core promoter element binding protein
	133838	M97786	BE222494	Hs.180919	inhibitor of DNA binding 2, dominant neg
	133859	U86782	U86782	Hs.178761	26S proteasome-associated pad1 homolog
	133889	AA099391	U48959	Hs.211582	myosin, light polypeptide kinase
	133960	M19267	M19267	Hs.77889	tropomyosin 1 (alpha)
15	133975	D29992	C18356	Hs.295944	tissue factor pathway inhibitor 2
	133977	L19314	AI125639	Hs.250666	hairy (Drosophila)-homolog
	134039	S78569	NM_002290	Hs.78672	tannin, alpha 4
	134075	U28811	NM_012201	Hs.78979	Golgi apparatus protein 1
	134081	L77886	AL034349	Hs.79005	protein tyrosine phosphatase, receptor t
20	134164	C14407	AW245540	Hs.79516	brain abundant, membrane attached signal
	134203	M60278	AA161219	Hs.799	diphtheria toxin receptor (haptin-bind)
	134238	R81509	AA102179	Hs.160726	Homo sapiens cDNA FLJ11680 fis, clone HE
	134299	AA487558	AW580939	Hs.97199	complement component C1q receptor
25	134332	D86962	D86962	Hs.81875	growth factor receptor-bound protein 10
	134339	AA478971	R70429	Hs.81888	disabled (Drosophila) homolog 2 (milogen
	134343	D50683	D50683	Hs.82028	transforming growth factor, beta recepto
	134381	U56637	AI557280	Hs.184270	capping protein (actin filament) muscle
	134403	M61199	AA334551		sperm specific antigen 2
30	134416	M28882	X68264	Hs.211579	melanoma cell adhesion molecule
	134493	X15183	M30627	Hs.289088	heat shock 90kD protein 1, alpha
	134558	SS3911	NM_001773	Hs.85289	CD34 antigen
	134617	U20734	AJ076592	Hs.198951	Jun B proto-oncogene
	134983	D28235	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p
35	134989	AA236324	AW968058	Hs.92381	nudix (nucleoside diphosphate linked moi
	135052	AA148823	AL136653	Hs.93675	deductal protein induced by progesterone
	135062	AA174183	AK000967	Hs.93872	KJAA1682 protein
	135069	AA456311	AA876372	Hs.93961	Homo sapiens mRNA; cDNA DKFZp667D085 (fr
	135071	L08069	W27190	Hs.94	DnaJ (Hsp40) homolog, subfamily A, membe
40	135073	AA452000	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp566E1624 (f
	135170	AA282140	T53169	Hs.9587	Homo sapiens cDNA: FLJ22290 fis, clone H
	135198	J02854	C03577	Hs.9615	myosin regulatory light chain 2, smooth
	135348	AA442054	U80983	Hs.268177	phospholipase C, gamma 1 (formerly subty
	134404	AB000450	AB000450	Hs.82771	vaccinia related kinase 2
45	439561	AB002380	AF180681	Hs.6582	Rho guanine exchange factor (GEF) 12
	100082	AB003103	AA130080	Hs.4295	proteasome (prosome, macropain) 26S subu
	132817	AB004884	N27852	Hs.57553	tousled-like kinase 2
	130150	AF000573	BE094848	Hs.15113	homogentisate 1,2-dioxygenase (homogenti
	100104	AF008937	AF008937		syntaxin 16
	447973	AF009301	AB011169	Hs.20141	similar to S. cerevisiae SSM4
50	332613	AF009368	AF029674	Hs.173422	KJAA1605 protein
	100113	D00591	NM_001269	Hs.84746	chromosome condensation 1
	133980	D00760	AA294921	Hs.348024	v-rat simian leukemia viral oncogene hom
	100129	D11139	AA469369	Hs.5831	tissue inhibitor of metalloproteinase 1
55	100154	D14657	H60720	Hs.81892	KJAA0101 gene product
	100169	D14878	AL037228	Hs.82043	D123 gene product
	129718	D17716	NM_002410	Hs.121502	mannosyl (alpha-1,6)-glycoprotein beta-
	100190	D21090	M91401	Hs.178658	RAD23 (S. cerevisiae) homolog B
	134742	D26135	NM_001346	Hs.89462	diacylglycerol kinase, gamma (90kD)
60	100211	D26528	D26528	Hs.123058	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
	100238	D30742	L24959	Hs.348	calcium/calmodulin-dependent protein kin
	130283	D31762	NM_012288	Hs.153954	TRAM-like protein
	134237	D31765	D31765	Hs.170114	KJAA0061 protein
	100248	D31888	NM_015156	Hs.78398	KJAA0071 protein
65	100256	D38128	D25418	Hs.393	prostaglandin I2 (prostaglandin) receptor
	100262	D38500	D38500	Hs.278468	postmeiotic segregation increased 2-like
	134329	D38551	N92036	Hs.81848	RAD21 (S. pombe) homolog
	100281	D42087	AF091035	Hs.184627	KJAA0118 protein
	100294	D49396	AA331881	Hs.75454	peroxiredoxin 3
70	100327	D55640	D55640		gbt human monocyte PABL (pseudautosomal
	100335	D63391	AW247529	Hs.6793	platelet-activating factor acetylhydrola
	134495	D63477	D63477	Hs.84087	KJAA0143 protein
	100338	D63483	D66864	Hs.57735	acetyl LDL receptor; SREC
	135152	D64015	M98954	Hs.182741	TIA1 cytotoxic granule-associated RNA-bi
75	134269	D79990	NM_014737	Hs.80905	Ras association (RalGDS/AF-6) domain fam
	100372	D79997	NM_014791	Hs.184339	KJAA0175 gene product
	134304	D80010	BE613486	Hs.81412	lipin 1

	100394	D84276	D84284	Hs.66052	CD38 antigen (p45)
	100405	D86425	AW291587	Hs.82733	nidogen 2
	100418	D86978	D86978	Hs.84790	KIAA0225 protein
	133154	D87012	D87012	Hs.194685	topoisomerase (DNA) III beta
5	134347	D87075	AF164142	Hs.82042	solute carrier family 23 (nucleobase tra
	444099	D87432	D87432	Hs.10315	solute carrier family 7 (cationic amino
	100438	D87448	AA013051	Hs.91417	topoisomerase (DNA) II binding protein
	134593	D87845	NM_000437	Hs.234392	platelet-activating factor acetylhydrola
	100481	HG1098-HT1098	X70377	Hs.121489	cystatin D
10	100552	HG2167-HT2237	AA019521	Hs.301946	lysosomal
	100591	HG2415-HT2511	NM_004091	Hs.231444	Homo sapiens, Similar to hypothetical pr
	100652	HG2825-HT2949	BE613608	Hs.142653	ret finger protein
	100662	HG2887-HT3031	AI368580	Hs.816	SRY (sex determining region Y)-box 2
15	100899	HG4660-HT5073	AL039123	Hs.103042	microtubule-associated protein 18
	100905	HG4704-HT5146	L12260	Hs.172816	neuregulin 1
	100945	HG884-HT884	AF002225	Hs.180688	ubiquitin protein ligase E3A (human papi
	100950	HG919-HT919	AF128542	Hs.166846	polymerase (DNA directed), epsilon
	100984	J00212	J00212		Empirically selected from AFTX single pr
20	135407	J04029	J04029	Hs.99936	keratin 10 (epidermolytic hyperkeratosis
	130149	J04031	AW067805	Hs.172665	methyleneetetrahydrofolate dehydrogenase
	131877	J04088	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)
	101016	J04543	J04543	Hs.78637	annexin A7
	134786	L06139	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous
	134100	L07540	AA460085	Hs.171075	replication factor C (activator 1) 5 (36
25	134078	L08895	L08895	Hs.78995	MADS box transcription enhancer factor 2
	101132	L11239	L11239	Hs.36993	gastrulation brain homeo box 1
	134849	L11353	BE409525	Hs.902	neurofibromin 2 (bilateral acoustic neur
	332736	L13773	ZB3689	Hs.114765	myeloid/lymphoid or mixed-lineage leukem
	101152	L13800	AI984625	Hs.9884	spindle pole body protein
30	135397	L14922	L14922	Hs.166563	replication factor C (activator 1) 1 (14
	432642	L15189	BE297635	Hs.3069	heat shock 70kD protein 9B (mortalin-2)
	101168	L15388	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	421155	L16895	H87879	Hs.102267	lysyl oxidase
35	101226	L27476	AF083892	Hs.75608	tight junction protein 2 (zona occludens
	415138	L27624	C18356	Hs.295944	tissue factor pathway inhibitor 2
	134739	L32976	NM_002419	Hs.89449	mitogen-activated protein kinase kinase
	130155	L33404	AA101043	Hs.151254	kallikrein 7 (chymotryptic, stratum com
	440538	L35263	W76332	Hs.79107	mitogen-activated protein kinase 14
	409916	L37347	BE313625	Hs.57435	solute carrier family 11 (proton-coupled
40	101294	L40371	AF168418	Hs.116784	thyroid hormone receptor interactor 4
	101300	L40391	BE535511		transmembrane trafficking protein
	101310	L41607	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2, I
	130344	L77566	AW250122	Hs.154879	DiGeorge syndrome critical region gene D
45	101381	M13928	AW675039	Hs.1227	aminolevulinic acid, delta-, dehydratase
	101381	M13928	AW675039	Hs.1227	aminolevulinic acid, delta-, dehydratase
	415678	M14016	AW005903	Hs.78601	uroporphyrinogen decarboxylase
	133780	M14219	AA557660	Hs.76152	decorin
	101396	M15796	BE267931	Hs.78996	proliferating cell nuclear antigen
50	101447	M21305	M21305		gbzHuman alpha satellite and satellite 3
	101458	M22092	M22092		gbzHuman neural cell adhesion molecule (
	101470	M22898	NM_000546	Hs.1846	tumor protein p53 (Li-Fraumeni syndrome)
	134604	M22995	NM_002884	Hs.865	RAP1A, member of RAS oncogene family
	101478	M23379	NM_002890	Hs.758	RAS p21 protein activator (GTPase activa
55	133519	M24400	AW583062	Hs.74502	chymotrypsinogen B1
	131185	M25753	BE280074	Hs.23960	cyclin B1
	134116	M27691	R84694	Hs.79194	cAMP responsive element binding protein
	133999	M28213	AA535244	Hs.78305	RAB2, member RAS oncogene family
	130174	M29550	M29551	Hs.151531	protein phosphatase 3 (formerly 2B), cat
	129963	M29971	M29971	Hs.1384	O-6-methylguanine-DNA methyltransferase
60	132983	M30269	M30269		nidogen (enactin)
	133900	M31158	M31158	Hs.77439	protein kinase, cAMP-dependent, regulato
	101543	M31166	M31166	Hs.2050	pentaxin-related gene, rapidly induced b
	101545	M31210	BE246154	Hs.154210	endothelial differentiation, sphingolip
	101620	M55420	S55271	Hs.247930	Epsilon, IgE
65	134691	M59979	AW382987	Hs.88474	prostaglandin-endoperoxide synthase 1 (p
	133595	M62810	AA393273	Hs.75133	transcription factor 6-like 1 (mitochond
	101700	M64710	D90337	Hs.247916	atriuretic peptide precursor C
	101714	M68874	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
70	134246	M74524	D28459	Hs.80612	ubiquitin-conjugating enzyme E2A (RAD6 h
	101760	M80254	M80254	Hs.173125	peptidylprolyl isomerase F (cyclophilin
	415022	M81780	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
	415022	M81780	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
	415022	M81780	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
	415022	M81780	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
75	415022	M81780	X59960	Hs.77813	sphingomyelin phosphodiesterase 1, acid
	101791	M83822	M83822	Hs.62354	cell division cycle 4-like

	101812	M86934	BE439894	Hs.78991	DNA segment, numerous copies, expressed
	101813	M87338	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40
	133396	M96326	M96326	Hs.72885	azurocidin 1 (cationic antimicrobial pro
	428161	M96954	M96954	Hs.182741	TIA1 cytotoxic granule-associated RNA-bi
5	129026	M98833	AL120297	Hs.108043	Friend leukemia virus integration 1
	101901	S66793	H38026	Hs.308	arrestin 3, retinal (X-arrestin)
	134831	S72370	AA853479	Hs.69890	pyruvate carboxylase
	134039	S78569	NM_002290	Hs.78672	laminin, alpha 4
	442355	S79873	AA456539	Hs.8262	lysosomal-associated membrane protein 2
10	101975	S83325	AA079717	Hs.283664	aspartate beta-hydroxylase
	101977	S83364	AF112213	Hs.184062	putative Rab5-interacting protein
	101978	S83365	BE561810	Hs.5809	putative transmembrane protein; homolog
	101998	U011212	U01212	Hs.248153	olfactory marker protein
	102003	U01922	U01922	Hs.125565	translocase of inner mitochondrial membr
15	102007	U02556	U02556	Hs.75307	t-complex-associated-testis-expressed 1-
	102009	U02680	BE245149	Hs.82643	protein tyrosine kinase 9
	416658	U03272	U03272	Hs.79432	fibrillin 2 (congenital contractural ara
	132951	U04209	AW821182	Hs.61418	microfibrillar-associated protein 1
	135389	U05237	U05237	Hs.99872	fetal Alzheimer antigen
20	102048	U07225	U07225	Hs.339	purinergic receptor P2Y, G-protein coupl
	130145	U07620	U34820	Hs.151051	mitogen-activated protein kinase 10
	303153	U09759	U09759	Hs.246857	mitogen-activated protein kinase 9
	420269	U09820	U72937	Hs.96264	alpha thalassemia/mental retardation syn
	102095	U11313	U11313	Hs.75760	sterol carrier protein 2
25	102123	U14518	NM_001809	Hs.1594	centromere protein A (17kD)
	102126	U14575	AW950870	Hs.78961	protein phosphatase 1, regulatory (inhib
	102133	U15173	AU076845	Hs.155596	BCL2/adenovirus E1B 19kD-interacting pro
	102139	U15932	NM_004419	Hs.2128	dual specificity phosphatase 5
	102162	U18291	AA450274	Hs.1592	CDC16 (cell division cycle 16, S. cerevi
30	102164	U18300	NM_000107	Hs.77602	damage-specific DNA binding protein 2 (4
	427653	U18383	AA159001	Hs.180069	nuclear respiratory factor 1
	131817	U20538	U20538	Hs.3280	caspase 6, apoptosis-related cysteine pr
	102200	U21551	AA232362	Hs.157205	branched chain aminotransferase 1, cytos
	102210	U23028	BE619413	Hs.2437	eukaryotic translation initiation factor
35	102214	U23752	U23752	Hs.32964	SRY (sex determining region Y)-box 11
	132811	U25435	U25435	Hs.57419	CCCTC-binding factor (zinc finger protei
	131319	U25997	NM_003155	Hs.25590	stanniocalcin 1
	102256	U28251	U28251	Hs.53237	ESTs, Highly similar to Z169_HUMAN ZINC
	132316	U28831	U28831	Hs.44566	KIAA1641 protein
40	102269	U30245	U30245		gldHuman myelomonocytic specific protein
	417526	U32315	AA568906	Hs.82240	syntaxin 3A
	102293	U32439	AF090116	Hs.79348	regulator of G-protein signalling 7
	102298	U32849	AA382169	Hs.54483	N-myc (and STAT) Interactor
	102325	U35139	AI815867	Hs.50130	neodlin (mouse) homolog
45	428734	U36764	BE303044	Hs.192023	eukaryotic translation initiation factor
	102361	U39400	AA223516	Hs.75859	chromosome 11 open reading frame 4
	102367	U39657	U39656	Hs.118825	mitogen-activated protein kinase kinase
	102388	U41344	AA362907	Hs.76494	proline arginine-rich end leucine-rich r
	102394	U41766	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma
50	129829	U41813	AF010258	Hs.127428	homeo box A9
	102409	U43286	BE300330	Hs.118725	selenophosphate synthetase 2
	133746	U44378	AW410035	Hs.75862	MAD (mothers against decapentaplegic, Dr
	102423	U44754	Z47542	Hs.179312	small nuclear RNA activating complex, po
	132828	U47011	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-ind
55	132828	U47011	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-ind
	132828	U47011	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-ind
	132828	U47011	AB014615	Hs.57710	fibroblast growth factor 8 (androgen-ind
	425322	U47077	U63630	Hs.155637	protein kinase, DNA-activated, catalytic
	102450	U48251	U48251	Hs.75871	protein kinase C binding protein 1
60	129350	U50535	U50535	Hs.110630	Human BRCA2 region, mRNA sequence CG006
	102534	U56833	U96759	Hs.198307	von Hippel-Lindau binding protein 1
	130457	U58091	AB014595	Hs.155976	cullin 4B
	135065	U58837	AA019401	Hs.93909	cyclic nucleotide gated channel beta 1
	102560	U59289	R97457	Hs.63984	cadherin 13, H-cadherin (heart)
65	102567	U59863	U63830	Hs.146847	TRAF family member-associated NFKB activ
	417173	U67122	U61397	Hs.81424	ubiquitin-like 1 (sentrin)
	102638	U67319	U67319	Hs.9216	caspase 7, apoptosis-related cysteine pr
	132736	U68019	AW081883	Hs.211578	Homo sapiens cDNA: FLJ23037 fls, clone L
	133070	U69611	U92649	Hs.64311	a disintegrin and metalloproteinase doma
70	102663	U70322	NM_002270	Hs.168075	karyopherin (importin) beta 2
	134660	U73524	U73524	Hs.87465	ATP/GTP-binding protein
	102735	U79267	AF111106	Hs.3382	protein phosphatase 4, regulatory subunl
	102741	U79291	AW959829	Hs.83572	hypothetical protein MGC14433
	130564	U82671	U82671	Hs.36980	melanoma antigen, family A, 2
75	130564	U82671	U82671	Hs.36980	melanoma antigen, family A, 2
	132164	U84573	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio

5	102823	U90914	D85390	Hs.5057	carboxypeptidase D
	102823	U91316	NM_007274	Hs.8879	cytosolic acyl coenzyme A thioester hyd
	102831	U91932	AA262170	Hs.80917	adaptor-related protein complex 3, sigma
	102846	U96131	BE264974	Hs.6566	thyroid hormone receptor interactor 13
	129777	U97018	U97018	Hs.12451	ectinoderm microtubule-associated protai
10	134161	U97188	AA634543	Hs.79440	IGF-II mRNA-binding protein 3
	134854	V00503	J03464	Hs.179573	collagen, type I, alpha 2
	429257	X04327	AW163789	Hs.198365	2,3-bisphosphoglycerate mutase
	413985	X06389	AJ018666	Hs.75667	synaptophysin
	419768	X07496	T72104	Hs.93194	apolipoprotein A-I
15	102915	X07820	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	134656	X14787	AJ750878	Hs.87409	thrombospondin 1
	413858	X15525	NM_001610	Hs.75589	acid phosphatase 2, lysosomal
	102968	X16396	AU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase
	102971	X16609	X16609	Hs.183805	ankyrin 1, erythrocytic
20	134037	X53586	AJ808780	Hs.227730	integrin, alpha 6
	134037	X53586	AJ808780	Hs.227730	integrin, alpha 6
	103023	X53793	AW500470	Hs.117950	multifunctional polypeptide similar to S
	103037	X54936	BE018302	Hs.2894	placental growth factor, vascular endoth
	130282	X55740	BE245380	Hs.153952	5' nucleotidase (CD73)
25	134542	X57025	M14156	Hs.85112	insulin-like growth factor 1 (somatomedi
	128568	X60673	H12912	Hs.274691	adenylate kinase 3
	128568	X60673	H12912	Hs.274691	adenylate kinase 3
	103093	X60708	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine
	413076	X62048	U10564	Hs.75188	wee1 (S. pombe) homolog
30	129063	X63097	X63094	Hs.283822	Rhesus blood group, D antigen
	424460	X63563	BE275979	Hs.296014	polymerase (RNA) II (DNA directed) polyp
	411077	X64037	AW977263	Hs.68257	general transcription factor IIF, polype
	103181	X69636	X69636	Hs.334731	Homo sapiens, clone IMAGE:3448306, mRNA,
	103184	X69878	U43143	Hs.74049	fms-related tyrosine kinase 4
35	103194	X70649	NM_004939	Hs.78580	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
	103208	X72841	AW411340	Hs.31314	retinoblastoma-binding protein 7
	129698	X74987	BE242144	Hs.12013	ATP-binding cassette, sub-family E (OABP
	131486	X83107	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	130729	X84194	AJ963747	Hs.18573	acylphosphatase 1, erythrocyte (common)
40	103334	X85753	NM_001260	Hs.25283	cyclin-dependent kinase 8
	132645	X87870	AJ654712	Hs.54424	hepatocyte nuclear factor 4, alpha
	135094	X89066	NM_003304	Hs.250687	transient receptor potential channel 1
	103352	X89398	H09366	Hs.78853	uracil-DNA glycosylase
	103352	X89398	H09366	Hs.78853	uracil-DNA glycosylase
45	103353	X89399	X89399	Hs.119274	RAS p21 protein activator (GTPase activa
	132173	X89426	X89426	Hs.41716	endothelial cell-specific molecule 1
	103371	X91247	X91247	Hs.13046	thioredoxin reductase 1
	131584	X91648	AA598509	Hs.29117	purine-rich element binding protein A
	103376	X92098	AL036166	Hs.323378	coated vesicle membrane protein
50	103378	X92110	AL119690	Hs.153618	HCGV11-1 protein
	128510	X94703	X94703		RAB28, member RAS oncogene family
	103410	X96506	AA158294	Hs.295362	DR1-associated protein 1 (negative cofac
	133490	X97230	AF022044	Hs.274601	killer cell immunoglobulin-like receptor
	132689	X97230	AF022044	Hs.274601	killer cell immunoglobulin-like receptor
55	103438	X98263	AW175781	Hs.152720	M-phase phosphoprotein 6
	103440	X98296	X98296	Hs.77578	ubiquitin specific protease 9, X chromos
	103452	X99584	NM_006936	Hs.85119	SMT3 (suppressor of mit two 3, yeast) ho
	133536	Y00264	W25797.comp		Hs.177486 amyloid beta (A4) precursor protein (pro
	420234	Y07566	AW404908	Hs.96038	Ric (Drosophila)-like, expressed in many
60	426502	Y07759	Y07759	Hs.170157	myosin VA (heavy polypeptide 12, myosin)
	134662	Y07827	NM_007048	Hs.284283	butyrophilin, subfamily 3, member A1
	132083	Y07867	BE386490	Hs.279663	Pitin
	103500	Y09443	AW408009	Hs.22580	alkylglycerone phosphate synthase
	134389	Y09858	Y09858	Hs.82577	spindlin-like
65	132084	Z112394	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)
	103540	Z11559	NM_002197	Hs.154721	aconitase 1, soluble
	133152	Z11695	Z11695	Hs.324473	mitogen-activated protein kinase 1
	103548	Z15005	Z15005	Hs.75573	centromere protein E (312kD)
	103612	Z46261	BE336654	Hs.70937	H3 histone family, member A
70	129092	AA011243	D58385	Hs.63525	poly(rC)-binding protein 2
	103692	AA018418	AW137912	Hs.227583	Homo sapiens chromosome X map Xp11.23 L-
	103695	AA018758	AW207152	Hs.186600	ESTs
	129796	AA018804	BE218319	Hs.5807	GTPase Rab14
	434993	AA031993	AA306325	Hs.4311	SUMO-1 activating enzyme subunit 2
75	132683	AA044217	BE264633	Hs.143638	WD repeat domain 4
	131887	AA046548	W17064	Hs.332848	SWI/SNF related, matrix associated, acti
	103723	AA057447	BE274312	Hs.214783	Homo sapiens cDNA FLJ14041 fs, clone HE
	453368	AA058376	W20296	Hs.288178	Homo sapiens cDNA FLJ11968 fs, clone HE
	133260	AA063572	AA403045	Hs.6906	Homo sapiens cDNA: FLJ2137 fs, clone R
	103765	AA085696	AA085696	Hs.169600	KIAA0826 protein

	103766	AA088744	AI920783	Hs.191435	ESTs
	103767	AA089688	BE244667		CGI-100 protein
	132051	AA091284	AA393968	Hs.180145	HSPC030 protein
5	103773	AA092700	AI219323	Hs.101077	ESTs, Weakly similar to T22363 hypothetical
	135289	AA092968	AW372569	Hs.9788	hypothetical protein MGC10924 similar to
	409659	AA094800	AW970843	Hs.55682	eukaryotic translation initiation factor
	103794	AA100219	AF244135	Hs.30870	hepatocellular carcinoma-associated anti
	131471	AA114885	AA164842	Hs.192619	KIAA1600 protein
10	134319	AA129547	BE304999	Hs.285754	fumarate hydratase
	103807	AA133016	AW958264	Hs.103832	similar to yeast UpK3, variant B
	446392	AA149507	AF142419	Hs.15020	homolog of mouse quaking QKI (KH domain
	129863	AA151005	BE379765	Hs.129872	sperm associated antigen 9
	103850	AA187101	AA187101	Hs.213194	hypothetical protein MGC10895
15	103855	AA195179	W02363		hypothetical protein FLJ10330
	103861	AA208236	AA208236	Hs.4944	hypothetical protein FLJ12783
	130634	AA227621	AI769087	Hs.127824	ESTs, Weakly similar to T28770 hypothetical
	447735	AA248283	AA775268	Hs.6127	Homo sapiens cDNA: FLJ23020 fs, clone L
	103909	AA249611	AA249611	Hs.47438	SH3 domain binding glutamic acid-rich pr
20	458928	AA282640	AF043117	Hs.24594	ubiquitination factor E4B (homologous to
	415824	AA287199	D42039	Hs.78871	mesoderm development candidate 2
	129013	AA313990	AA371156	Hs.107942	DKFZP564M112 protein
	129435	AA314256	AF151852	Hs.111449	CGI-94 protein
	103988	AA314389	AA314389	Hs.342849	ADP-ribosylation factor-like 5
25	104000	AA324364	AI146527	Hs.80475	polymerase (RNA) II (DNA directed) polyp
	425284	AA329211	AF155568	Hs.348043	NS1-associated protein 1
	128629	AA399187	AL096748	Hs.102708	DKFZP434A043 protein
	133281	AA421079	AK001601	Hs.69594	high-mobility group 20A
	104104	AA422029	AA422029	Hs.143640	ESTs, Weakly similar to hyperpolarizatio
30	332455	AA425230	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain
	132091	AA447052	AW954243		KIAA0251 protein
	135073	AA452000	W55958	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
	131367	AA456687	AI750575	Hs.173933	nuclear factor I/A
35	129593	AA487015	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f
	133505	C011527	AI630124	Hs.324504	Homo sapiens mRNA; cDNA DKFZp586J0720 (f
	132064	C01714	AA121098	Hs.3838	serum-inducible kinase
	442351	C01811	W52642	Hs.8261	hypothetical protein FLJ22393
	131427	C02352	AF151879	Hs.26708	CGI-121 protein
40	433892	C02375	AI929357	Hs.323966	Homo sapiens clone H63 unknown mRNA
	104282	C14448	C14448	Hs.332338	EST
	134827	D16611	BE314037	Hs.89866	coproporphyrinogen oxidase (coproporphyr
	425330	D25216	D25216	Hs.155650	KIAA0014 gene product
	131742	D31352	AA961420	Hs.31433	ESTs
	456935	D58024	AA370362	Hs.57958	EGF-TM7-latrophilin-related protein
45	425218	D80897	NM_014909	Hs.155182	KIAA1036 protein
	104334	D82614	D82614	Hs.78771	phosphoglycerate kinase 1
	134593	D87845	NM_000437	Hs.234392	platelet-activating factor acetylhydrola
	134731	D89377	D89377	Hs.89404	msH (Drosophila) homeo box homolog 2
	445776	H06583	NM_001310	Hs.13313	cAMP responsive element binding protein-
50	131670	H40732	H03514	Hs.15589	ESTs
	104394	H46617	AA129551	Hs.172129	Homo sapiens cDNA: FLJ21409 fs, clone C
	104402	H56731	H56731	Hs.132956	ESTs
	439130	H75570	AA306090	Hs.124707	ESTs
	129077	H78886	N74724	Hs.108479	ESTs
55	104417	H81241	AI819448	Hs.320861	Kruppel-like factor 8
	134927	L36531	L36531	Hs.91296	integrin, alpha 8
	129280	M63154	M63154	Hs.110014	gastric intrinsic factor (vitamin B synt
	134498	M63180	AW246273	Hs.84131	threonyl-tRNA synthetase
	104460	M91504	AW955705	Hs.62604	Homo sapiens, clone IMAGE:4299322, mRNA,
60	104488	N56191	N56191	Hs.106511	protocadherin 17
	131248	N78483	AI038989	Hs.332633	Bardet-Biedl syndrome 2
	130017	R14652	AK000096	Hs.143198	inhibitor of growth family, member 3
	104530	R20459	AK001676	Hs.12457	hypothetical protein FLJ10814
	104534	R22303	R22303		gbbyh26b09.r1 Soares placenta Nb2HP Homo
65	104544	R33779	AI091173	Hs.222362	ESTs, Weakly similar to p40 [Hsapiens]
	133328	R36553	AW452738	Hs.265327	hypothetical protein DKFZp761141
	104567	R64534	AA040620	Hs.5672	hypothetical protein AF140225
	129575	R70621	F08282	Hs.278428	progesterone induced protein
	130776	R79356	AF167708	Hs.19280	cysteine-rich motor neuron 1
70	104599	R84933	AW815036	Hs.151251	ESTs
	104660	AA007150	BE296665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr
	104667	AA007234	AI239923	Hs.63931	ESTs
	104718	AA018409	AI143020	Hs.36250	ESTs, Weakly similar to I38022 hypotheti
	104764	AA025351	AI039243	Hs.278585	ESTs
75	104786	AA027168	AA027167	Hs.10031	KIAA0955 protein
	104787	AA027317	AA027317		gbcze7d11.s1 Soares_fetal_heart_NbIH19W
	134079	AA029423	AK001751	Hs.171835	hypothetical protein FLJ10889

	104804	AA031357	AJ858702	Hs.31803	ESTs, Weakly similar to N-WASP [Hsapien
	104865	AA045136	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
	130828	AA053400	AW631469	Hs.203213	ESTs
5	104907	AA055829	AA055829	Hs.198701	ESTs, Weakly similar to ALU1_HUMAN ALU S
	104943	AA065217	AF072873	Hs.114218	frizzled (Drosophila) homolog 6
	105013	AA116054	H63789	Hs.296288	ESTs, Weakly similar to KIAA0638 protein
	105024	AA126311	AA126311	Hs.9879	ESTs
	132582	AA129390	AW803564	Hs.288650	Homo sapiens cDNA: FLJ22528 fis, clone H
10	105038	AA130273	AW503733	Hs.9414	KIAA1488 protein
	105077	AA142919	W55946	Hs.234863	Homo sapiens cDNA FLJ12082 fis, clone HE
	105096	AA150205	AL042506	Hs.21599	Kruppel-like factor 7 (ubiquitous)
	129215	AA176867	AB040930	Hs.126085	KIAA1497 protein
	105169	AA180321	BE245294	Hs.180789	S164 protein
15	132786	AA180487	NM_006283	Hs.173159	transforming, acidic coiled-coil contain
	427210	AA187634	BE396283	Hs.173987	eukaryotic translation initiation factor
	105200	AA195399	AA328102	Hs.24641	cytoskeleton associated protein 2
	130114	AA234717	AA233393	Hs.14992	hypothetical protein FLJ11151
	105330	AA234743	AW338625	Hs.22120	ESTs
20	105337	AA234957	AA665789	Hs.347187	myotubularin related protein 1
	422040	AA235604	AA172106	Hs.110950	Rag C protein
	105376	AA235659	AW994032	Hs.8768	hypothetical protein FLJ10849
	105397	AA242868	AA814807	Hs.7395	hypothetical protein FLJ23182
	431679	AA251776	AK000046	Hs.343877	hypothetical protein FLJ20039
25	131991	AA251909	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1
	421305	AA252672	BE397354	Hs.324830	diphtheria toxin resistance protein requi
	105489	AA256157	AA256157	Hs.24115	Homo sapiens cDNA FLJ14178 fis, clone NT
	105508	AA256680	AA173942	Hs.326416	Homo sapiens mRNA; cDNA DKFZp664H1916 (f
	105539	AA258873	AB040884	Hs.109694	KIAA1451 protein
30	135172	AA262727	AB028956	Hs.12144	KIAA1033 protein
	131569	AA281451	AL389951	Hs.271623	nucleoporin 50kD
	431129	AA281545	AL137751	Hs.263671	Homo sapiens mRNA; cDNA DKFZp434I0812 (f
	105643	AA282069	BE621719	Hs.173802	KIAA0603 gene product
	105659	AA283044	AA283044	Hs.25625	hypothetical protein FLJ11323
35	105666	AA283930	AA426234	Hs.34906	ESTs, Weakly similar to T17210 hypotheti
	105674	AA284755	AI609530	Hs.279789	histone deacetylase 3
	105709	AA291268	AI928962	Hs.26761	DKFZP586L0724 protein
	105722	AA291927	AI922821	Hs.32433	ESTs
	105765	AA343514	AA298668	Hs.24183	ESTs
40	115951	AA398109	BE546245	Hs.301048	sec13-like protein
	130884	AA398109	BE546245	Hs.301048	sec13-like protein
	105962	AA405737	AW880358	Hs.339808	hypothetical protein FLJ10120
	105985	AA406610	AA406610		gbzv15b10.s1 Soares_NHMPu_S1 Homo sapi
	106008	AA411465	AB033888	Hs.8619	SRY (sex determining region Y)-box 18
45	457322	AA416886	AI815486	Hs.243901	Homo sapiens cDNA FLJ20738 fis, clone HE
	134222	AA424013	AW855861	Hs.8025	Homo sapiens clone 23767 and 23782 mRNA
	448954	AA424148	AB037850	Hs.16621	DKFZP434I116 protein
	106141	AA424558	AF031463	Hs.9302	phosducin-like
	447973	AA424961	AB011169	Hs.20141	similar to S. cerevisiae SSM4
50	106157	AA425367	W37943	Hs.34892	KIAA1323 protein
	428314	AA425921	AW135049	Hs.26285	Homo sapiens cDNA FLJ10643 fis, clone NT
	446727	AA426220	AB011095	Hs.18032	KIAA0523 protein
	106196	AA427735	AA525993	Hs.173699	ESTs, Weakly similar to ALU1_HUMAN ALU S
	457714	AA430673	AA083764		hypothetical protein MGC3178
55	133200	AA432248	AB037715	Hs.183639	hypothetical protein FLJ10210
	106302	AA435896	AA398859	Hs.18397	hypothetical protein FLJ23221
	106328	AA436705	AL079559	Hs.28020	KIAA0766 gene product
	450534	AA446561	AI570189	Hs.25132	KIAA0470 gene product
	106423	AA448238	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15
60	439608	AA449756	AW864696	Hs.301732	hypothetical protein MGC5306
	106477	AA450303	R23324	Hs.41693	DnaJ (Hsp40) homolog, subfamily B, membe
	106503	AA452411	AB033042	Hs.29679	cofactor required for Sp1 transcription
	446999	AA454566	AA151520		hypothetical protein MGC4485
	106543	AA454667	AA676939	Hs.69285	neuroplatin 1
	442007	AA456437	AA301118	Hs.142838	nucleolar phosphoprotein Nopp34
65	106589	AA456646	AK000933	Hs.28661	Homo sapiens cDNA FLJ10071 fis, clone HE
	106593	AA456826	AW296451	Hs.24605	ESTs
	106596	AA456981	AA452379		ESTs, Moderately similar to ALU7_HUMAN A
	423064	AA458959	AF265208	Hs.8740	SWI5NF related, matrix associated, acti
70	106636	AA459950	AW958037	Hs.286	ribosomal protein L4
	106654	AA460449	AW075485	Hs.286049	phosphoserine aminotransferase
	131353	AA463910	AW754182		gbRC2-CT0321-131199-011-c01 CT0321 Homo
	106707	AA464603	AK000566	Hs.98135	hypothetical protein FLJ20559
	452909	AA464606	NM_015368	Hs.30985	pannexin 1
	106717	AA465093	AA600357	Hs.239489	TIA1 cytotoxic granule-associated RNA-bl
75	453141	AA465692	AB014548	Hs.31921	KIAA0648 protein
	106747	AA476473	NM_007116	Hs.171957	triple functional domain (PTPRF) interact

	106773	AA478109	AA478109	Hs.188833	ESTs
	106781	AA478474	AA330310	Hs.24181	ESTs
	106817	AA480889	D61216	Hs.18672	ESTs
	106846	AA485223	AB037744	Hs.34892	KIAA1323 protein
5	106848	AA485254	AA449014	Hs.121025	chromosome 11 open reading frame 5
	106856	AA486183	W58353	Hs.285123	Homo sapiens mRNA full length insert cDN
	418639	AA496936	BE539639	Hs.173030	ESTs, Weakly similar to ALU8_HUMAN ALU S
	107001	AA588589	AJ926520	Hs.31016	putative DNA binding protein
	442853	AA588831	AW021276	Hs.17121	ESTs
10	107064	AA600150	AJ076459	Hs.15978	KIAA1272 protein
	107069	AA608545	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E. coli Re
	107080	AA609210	AL122043	Hs.19221	hypothetical protein DKFZp566G1424
	107115	AA610108	BE379623	Hs.27693	peptidylprolyl isomerase (cyclophilin)-I
	107130	AA620582	AB033106	Hs.12913	KIAA1280 protein
15	107156	AA621239	AA137043	Hs.9663	programmed cell death 6-interacting prot
	107174	AA621714	BE122762	Hs.25338	ESTs
	130621	AA621718	AW513087	Hs.16803	LJC7 (S. cerevisiae)-like
	107190	D19873	AA836401	Hs.87860	ESTs
	132626	D25755	AW504732	Hs.21275	hypothetical protein FLJ11011
20	107217	D51095	AL080235	Hs.35861	DKFZP586E1621 protein
	332584	D60272	AA357879	Hs.29423	ESTs; Weakly similar to macrophage lect
	444655	T08879	AF088886	Hs.11590	cathepsin F
	107295	T34527	AA186629	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
	107299	T40327	BE277457	Hs.30661	hypothetical protein MGC4606
25	107315	T62771	AA316241	Hs.90691	nucleophosmin/nucleoplasm 3
	107316	T63174	T63174	Hs.193700	Homo sapiens mRNA: cDNA DKFZp586I0324 (f
	107328	T83444	AW958891	Hs.76591	KIAA0687 protein
	107334	T93841	T93597	Hs.187429	ESTs
	456340	U48263	U48263	Hs.89040	prepronociceptin
30	128636	U49065	U49065	Hs.102865	interleukin 1 receptor-like 2
	129938	U79300	AW003668	Hs.135587	Human clone 23629 mRNA sequence
	107375	U88573	BE011845	Hs.251064	high-mobility group (nonhistone chromoso
	130074	U93867	AL038596	Hs.250745	polymerase (RNA) III (DNA directed) (62k
35	107387	W01094	D88983	Hs.118893	Melanoma associated gene
	132036	W01568	AL157433	Hs.37706	hypothetical protein DKFZp434E2220
	107426	W26853	W26853	Hs.291003	hypothetical protein MGC4707
	135388	W27965	W27965	Hs.98865	epimorphin
	130419	W36280	AF037448	Hs.155489	NS1-associated protein 1
	107469	W47063	W47063	Hs.94668	ESTs
40	434203	W79060	BE262677	Hs.283558	hypothetical protein PRO1855
	107506	W88550	AB028981	Hs.8021	KIAA1058 protein
	132358	X60486	NM_003542	Hs.46423	H4 histone family, member G
	107522	X78931	X78931	Hs.99971	zinc finger protein 272
	456495	Z14077	NM_003403	Hs.97496	YY1 transcription factor
45	107582	AA002147	AA002147	Hs.58952	EST
	107609	AA004711	R75654	Hs.164797	hypothetical protein FLJ13693
	107661	AA010383	AA010383	Hs.60389	ESTs
	107714	AA015761	AA015761	Hs.60642	ESTs
	107775	AA016772	AW008846	Hs.60857	ESTs
50	107832	AA021473	AA021473	Hs.61555	gb:z666c11.s1 Soares retina N2b4HR Homo
	107859	AA024835	AW732573	Hs.47584	potassium voltage-gated channel, delayed
	107914	AA027229	AA027229	Hs.61329	ESTs, Weakly similar to T16370 hypotheti
	107935	AA029428	AA029428	Hs.61555	ESTs
	410196	AA035143	AJ936442	Hs.59838	hypothetical protein FLJ10808
55	131461	AA035237	AA992841	Hs.27263	KIAA1458 protein
	108007	AA039347	AA039347	Hs.61916	EST
	108029	AA040740	AA040740	Hs.62007	ESTs
	108040	AA041551	AL121031	Hs.159971	SWI/SNF related, matrix associated, acti
	108084	AA045513	AA058944	Hs.116602	Homo sapiens, clone IMAGE:4154008, mRNA,
60	108088	AA045745	AA045745	Hs.62886	ESTs
	108168	AA055348	AJ453137	Hs.63176	ESTs
	130719	AA055582	AA679262	Hs.14235	hypothetical protein FLJ20008; KIAA1839
	108189	AA056697	AW376061	Hs.63335	ESTs, Moderately similar to A46010 X-fin
	108190	AA056746	AA056746	Hs.63338	EST
65	108203	AA057678	AW847814	Hs.289005	Homo sapiens cDNA: FLJ21532 fis, clone C
	108216	AA058681	AA524743	Hs.44883	ESTs
	108217	AA058686	AA058686	Hs.62588	ESTs
	108245	AA062840	BE410285	Hs.89545	proteasome (prosome, macropain) subunit,
	108277	AA064859	AA064859	Hs.89545	gb:zm50f03.s1 Stratagene fibroblast (937
70	108280	AA065069	AA065069	Hs.89545	gb:zm12e11.s1 Stratagene pancreas (93720
	108309	AA069923	AA069918	Hs.180909	gb:zm57e03.r1 Stratagene neuroepithelium
	108340	AA070815	AA069820	Hs.180909	peroxiredoxin 1
	108403	AA075374	AA075374	Hs.180909	gb:zm87a01.s1 Stratagene ovarian cancer
	108427	AA076382	AA076382	Hs.180909	gb:zm91g08.s1 Stratagene ovarian cancer
75	108435	AA078767	T82427	Hs.194101	Homo sapiens cDNA: FLJ20869 fis, clone A
	108439	AA078986	AA078986	Hs.194101	gb:zm92h01.s1 Stratagene ovarian cancer

	108465	AA079393	AA079393	Hs.3462	cytochrome c oxidase subunit VIc
	108469	AA079487	AA079487		gb:zm97108.s1 Stratagene colon HT29 (937
	108500	AA083207	AA083207	Hs.68270	EST
	108501	AA083256	AA083256		gb:zm08g12.s1 Stratagene hNT neuron (937
5	108533	AA084415	AA084415		gb:zm06g09.s1 Stratagene hNT neuron (937
	108562	AA085274	AA100796		gb:zm26c06.s1 Stratagene pancreas (93720
	108589	AA088678	A1732404	Hs.68846	ESTs
	130890	AA100925	A1907537	Hs.76698	stress-associated endoplasmic reticulum
	432845	AA101255	D14041	Hs.347340	H-2K binding factor-2
10	130385	AA126474	AW067800	Hs.155223	stannocalcin 2
	108749	AA127017	AA127017	Hs.71052	ESTs
	108807	AA129968	A1652236	Hs.49376	hypothetical protein FLJ20644
	108808	AA130240	AA045088	Hs.62738	ESTs
	108833	AA131866	AF188527	Hs.61681	ESTs, Weakly similar to AF174605 1 F-box
15	108846	AA132983	AL117452	Hs.44155	DKFZP586G1517 protein
	108857	AA133250	AK001458	Hs.62180	anillin (Drosophila Scraps homolog), act
	131474	AA133583	L48353	Hs.2726	high-mobility group (nonhistone chromoso
	108894	AA135941	AK001431	Hs.5105	hypothetical protein FLJ10569
	108941	AA148650	AA148650		gb:zm09e06.s1 Stratagene neuroepithelium
20	108968	AA151110	A1304870	Hs.188680	ESTs
	108996	AA155754	AW995610	Hs.332436	EST
	109001	AA156125	A1056548	Hs.72116	hypothetical protein FLJ20992 similar to
	131183	AA156289	A1611807	Hs.285107	hypothetical protein FLJ13397
	109019	AA156997	AA156755	Hs.72150	ESTs
25	109022	AA157291	AA157291	Hs.21479	ubiquitin 1
	109023	AA157293	AA157293	Hs.72168	ESTs
	109068	AA164293	AA164293	Hs.72545	ESTs
	109072	AA164676	A1732585	Hs.22394	hypothetical protein FLJ10893
	426981	AA167375	AL044675	Hs.173081	KIAA0530 protein
30	130346	AA167550	H05769	Hs.188757	Homo sapiens, clone MGC:5564, mRNA, comp
	109146	AA176589	AA176589	Hs.142078	EST
	109172	AA180448	AA180448	Hs.144300	EST
	428438	AA187144	NM_001955	Hs.2271	endothelin 1
	129208	AA189170	A1587376	Hs.109441	MSTP033 protein
35	109222	AA192757	AA192833	Hs.333512	similar to rat myomesin
	109300	AA205850	AA418276	Hs.170142	ESTs
	109481	AA233342	AA878923	Hs.289069	hypothetical protein FLJ21016
	109485	AA233472	BE619082	Hs.28465	Homo sapiens cDNA: FLJ21869 fs, clone H
40	109516	AA234110	A1471639	Hs.71913	ESTs
	109537	D80981	A1858695	Hs.34898	ESTs
	109556	F01660	A1925294	Hs.87385	ESTs
	109577	F02206	F02206	Hs.296639	Homo sapiens potassium channel subunit (
	109578	F02208	F02208	Hs.27214	ESTs
	109595	F02544	AA078629	Hs.27301	ESTs
45	109625	F03918	H29490	Hs.22697	ESTs
	428376	F04258	AF119665	Hs.184011	pyrophosphatase (inorganic)
	109648	F04600	H17800	Hs.7154	ESTs
	109671	F08998	R59210	Hs.26634	ESTs
	109699	F09605	H18013	Hs.167463	ESTs
50	109820	F11115	AW016809	Hs.119021	ESTs
	109933	H06371	R52417	Hs.20945	Homo sapiens clone 24993 mRNA sequence
	110014	H10995	AL109666	Hs.7242	Homo sapiens mRNA full length insert cDN
	110039	H11938	H11938	Hs.21907	histone acetyltransferase
	110099	H16568	R44557	Hs.23748	ESTs
55	110107	H16772	AW151660	Hs.31444	ESTs
	110155	H18951	A1559626	Hs.93522	Homo sapiens mRNA for KIAA1647 protein,
	110197	H20859	AW090386	Hs.112278	arrestin, beta 1
	110223	H23747	H19836	Hs.31697	ESTs
	110306	H38087	H38087	Hs.105509	CTL2 gene
60	110335	H40331	H65490	Hs.18845	ESTs
	110342	H40567	H40961	Hs.33008	ESTs
	110395	H46966	AA025116	Hs.33333	ESTs
	110511	H56640	H56640	Hs.221460	ESTs
	110523	H57154	A1040384	Hs.19102	ESTs, Weakly similar to organic anion tr
65	110715	H98712	H98712	Hs.269029	ESTs
	110754	N20814	AW302200	Hs.6336	KIAA0672 gene product
	428454	N25249	U55936	Hs.184376	synaptosomal-associated protein, 23kD
	431663	N27100	NM_016569	Hs.267182	TBX3-iso protein
	134263	N39616	AW973443	Hs.8086	RNA (guanine-7-) methyltransferase
70	110938	N48962	N48962	Hs.38034	Homo sapiens cDNA FLJ12924 fs, clone NT
	110983	N51857	NM_015367	Hs.10267	MIL1 protein
	111081	N59435	A146349	Hs.271614	CGI-112 protein
	111128	N64139	AW505364	Hs.19074	LATS (large tumor suppressor, Drosophila
	431548	N66981	A1834273	Hs.9711	novel protein
75	111216	N68640	AW139408	Hs.152940	ESTs
	437562	N69352	AB001636	Hs.5683	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep

	111399	R00138	AW270776	Hs.18857	ESTs
	111514	R07998	R07998		gb:Yf16g11.s1 Soares fetal liver spleen
	428744	R08929	BE267033	Hs.192853	ubiquitin-conjugating enzyme E2G 2 (homo
5	111574	R10307	AJ024145	Hs.188526	ESTs
	111804	R33354	AA482478	Hs.181785	ESTs
	111831	R36083	R36095	Hs.268695	ESTs
	426773	R37938	NM_015556	Hs.172180	KIAA0440 protein
	111904	R39330	Z41572		gb:HSCZYB122 normalized infant brain cDN
10	428371	R40816	AB012193	Hs.183674	outlin 4A
	112033	R43162	R49031	Hs.22627	ESTs
	130987	R45699	BE613269	Hs.21893	hypothetical protein DKFZp761N0624
	112300	R54554	H24334	Hs.26125	ESTs
	112513	R68425	R68425	Hs.13809	hypothetical protein FLJ10648
15	112514	R68568	R68568	Hs.183373	src homology 3 domain-containing protein
	112522	R68763	R68857	Hs.265499	ESTs
	112540	R70467	R69751		gb:Y40a10.s1 Soares placenta Nb2HP Homo
	428655	R73565	H05769	Hs.188757	Homo sapiens, clone MGC:5564, mRNA, comp
	129534	R73640	AK002126	Hs.11260	hypothetical protein FLJ11264
20	112597	R78376	R78376	Hs.29733	EST
	112732	R92453	R92453	Hs.34590	ESTs
	451798	T03865	BE297567	Hs.27047	hypothetical protein FLJ20392
	112888	T03872	AW195317	Hs.107716	hypothetical protein FLJ22344
	131863	T10072	AJ656378	Hs.33461	ESTs
25	112911	T10080	AW732747	Hs.13493	like mouse brain protein E46
	132215	T10132	AL035703	Hs.4236	KIAA0478 gene product
	112931	T15343	T02966	Hs.187428	ESTs
	112984	T23457	T16971	Hs.289014	ESTs, Weakly similar to A43932 mucin 2 p
	112998	T23555	H11257	Hs.22968	Homo sapiens clone IMAGE:451939, mRNA se
30	133376	T23670	BE618768	Hs.7232	acetyl-Coenzyme A carboxylase alpha
	113026	T23948	AA376654		eukaryotic translation initiation factor
	113070	T33464	AB032977	Hs.6298	KIAA1151 protein
	410781	T34413	AJ375672	Hs.165028	ESTs
	113074	T34611	AK001335	Hs.31137	protein tyrosine phosphatase, receptor t
35	113095	T40820	AA828380	Hs.126733	ESTs
	113179	T55182	BE622021	Hs.152571	ESTs, Highly similar to IGF-II mRNA-bind
	113337	T77453	T77453	Hs.302234	ESTs
	113421	T84039	AJ769400	Hs.189729	ESTs
	113454	T86458	AJ022166	Hs.16188	ESTs
40	113481	T87683	T87683	Hs.204327	EST
	453345	T89350	AA302862	Hs.90063	neurocalcin delta
	113557	T90945	H66470	Hs.18004	ESTs
	113559	T90987	T79763	Hs.14514	ESTs
	113589	T91863	AJ078554	Hs.15682	ESTs
45	113591	T91881	T91881	Hs.200597	KIAA0563 gene product
	113619	T93783	R08665	Hs.17244	hypothetical protein FLJ13605
	113683	T96687	AB035335	Hs.144519	T-cell leukemia/lymphoma 6
	113692	T96944	AL360143	Hs.17936	DKFZP434H132 protein
	113702	T97307	T97307		gb:ye53h05.s1 Soares fetal liver spleen
50	113717	T97764	T99513	Hs.187447	ESTs
	113824	W48817	AJ631964	Hs.34447	ESTs
	113840	W58343	R72137	Hs.7949	DKFZP586B2420 protein
	113844	W59949	AJ369275	Hs.243010	Homo sapiens cDNA FLJ14445 fls, clone HE
	113902	W74644	AA340111	Hs.100009	acyl-Coenzyme A oxidase 1, palmitoyl
55	113904	W74761	AF125044	Hs.19196	ubiquitin-conjugating enzyme HBUCE1
	113905	W74802	R81733	Hs.33106	ESTs
	113931	W81205	BE255499	Hs.3496	hypothetical protein MGC15749
	113932	W81237	AA256444	Hs.126485	hypothetical protein FLJ12604; KIAA1692
	131965	W90146	W78283	Hs.35962	ESTs
60	114035	W92798	W92798	Hs.269181	ESTs
	114106	Z38412	AW602528		gb:RC5-BT0562-260100-011-A02 BT0562 Homo
	457308	Z38709	AJ169888	Hs.238272	inositol 1,4,5-trisphosphate receptor, ty
	114161	Z38904	BE548222	Hs.299883	hypothetical protein FLJ23399
	424949	Z39103	AF052212	Hs.153934	core-binding factor, runt domain, alpha
	457548	Z39930	AW069534	Hs.279583	CGI-81 protein
65	128937	Z39939	AA251380	Hs.10726	ESTs, Weakly similar to ALU1_HUMAN ALU S
	432554	Z40012	AJ479813	Hs.278411	NCK-associated protein 1
	114277	Z40377	AJ052229	Hs.25373	ESTs, Weakly similar to T20410 hypotheti
	114304	Z40820	AJ934204	Hs.16129	ESTs
	114364	Z41680	AL117427	Hs.172778	Homo sapiens mRNA; cDNA DKFZp566P013 (fr
70	432620	AA005112	AA777749	Hs.5978	LIM domain only 7
	129034	AA005432	AA481157	Hs.108110	DKFZP547E2110 protein
	131881	AA010163	AW361018	Hs.3383	upstream regulatory element binding prot
	332421	AA026356	AJ909968	Hs.108106	transcription factor
	114465	AA026901	BE621056	Hs.131731	hypothetical protein FLJ11099
75	451271	AA036867	AK001644	Hs.28156	hypothetical protein FLJ10782
	332498	AA044644	AA303661		lymphocyte-specific protein 1

	431555	AA046426	AI815470	Hs.260024	Cdc42 effector protein 3
	132944	AA054515	T96641	Hs.6127	Homo sapiens cDNA: FLJ23020 fis, clone L
	114618	AA084162	AW979261	Hs.291993	ESTs
	332509	AA085749	AA128376	Hs.153884	ATP binding protein associated with cell
5	114648	AA101056	AA101056		gb:zn25b03.s1 Stratagene neuroepithelium
	114658	AA102746	AA102383	Hs.249190	tumor necrosis factor receptor superfamily
	132456	AA114250	AB011084	Hs.488924	KIAA0512 gene product; ALEX2
	450847	AA126561	NM_003155	Hs.255590	.stanniocalcin 1
	132225	AA128980	AA128980		gb:zo09a11.s1 Stratagene neuroepithelium
10	437197	AA129757	W38586		guanine nucleotide binding protein (G pr
	114709	AA129921	AA397651	Hs.301959	proline synthetase co-transcribed (bacte
	456926	AA133331	AB018284	Hs.158688	KIAA0741 gene product
	114750	AA135958	AA887211	Hs.129467	ESTs
	426806	AA136524	T19228	Hs.172572	hypothetical protein FLJ20093
15	114763	AA147044	AA810755	Hs.102500	hypothetical protein dJ511E16.2
	114767	AA148885	AB598665	Hs.154443	minichromosome maintenance deficient (S.
	114774	AA150043	AV656017	Hs.184325	CGI-76 protein
	128388	AA151621	AA662477	Hs.110964	hypothetical protein FLJ23471
	457742	AA155743	BE561824	Hs.273369	uncharacterized hematopoietic stem/proge
20	456200	AA156335	AA768242	Hs.80618	hypothetical protein
	130207	AA156336	AF044209	Hs.144904	nuclear receptor co-repressor 1
	114788	AA159181	AA159181	Hs.54900	serologically defined colon cancer antig
	114800	AA159825	Z19448	Hs.131887	ESTs, Weakly similar to T24396 hypotheti
	114828	AA234185	AA252937	Hs.283522	Homo sapiens mRNA; cDNA DKFZp434J1912 (f
25	114846	AA234929	BE018682	Hs.166196	ATPase, Class I, type 8B, member 1
	114848	AA234935	BE614347	Hs.169615	hypothetical protein FLJ20989
	114902	AA236359	AW275480	Hs.39504	hypothetical protein MGC4308
	132271	AA236466	AB030034	Hs.115175	sterile-alpha motif and leucine zipper c
	114907	AA236535	N29390	Hs.13804	hypothetical protein dJ462023.2
30	420170	AA236935	U43374	Hs.95631	Human normal keratinocyte mRNA
	132204	AA236942	AA235827	Hs.42265	ESTs
	114928	AA237018	AA237018	Hs.94869	ESTs
	132481	AA237025	W93378	Hs.49614	ESTs
	114932	AA242751	AA971436	Hs.16218	KIAA0903 protein
35	314162	AA242760	BE041820	Hs.38516	Homo sapiens, clone MGC:15887, mRNA, com
	131006	AA242763	AF064104	Hs.22116	CDC14 (cell division cycle 14, S. cerevis
	114935	AA242809	H23329	Hs.290680	ESTs, Weakly similar to ALU1_HUMAN ALU S
	408908	AA243133	BE296227	Hs.250822	serine/threonine kinase 15
	437754	AA243495	R60366	Hs.5822	Homo sapiens cDNA: FLJ22120 fis, clone H
40	114957	AA243706	AW170425	Hs.87680	ESTs
	114974	AA250848	AW566931	Hs.302649	nucleosome assembly protein 1-like 1
	114977	AA250868	AW296978	Hs.87787	ESTs
	114995	AA251152	AA769266	Hs.193657	ESTs
	115005	AA251544	AI760825	Hs.153042	ESTs
45	417177	AA251792	NM_004458	Hs.81452	fatty-acid-Coenzyme A ligase, long-chain
	115026	AA252144	AA251972	Hs.188718	ESTs
	115045	AA252524	AW014549	Hs.58373	ESTs
	115068	AA253461	AW512260	Hs.87767	ESTs
	133138	AA255522	AV657594	Hs.181161	Homo sapiens cDNA FLJ14643 fis, clone NT
50	332668	AA255522	AV657594	Hs.181161	ESTs
	115114	AA256468	AA527548	Hs.7527	small fragment nuclease
	129584	AA256528	AV656017	Hs.184325	CGI-76 protein
	115137	AA257976	AW968304	Hs.56156	ESTs
	417187	AA258296	AB011151	Hs.334659	hypothetical protein MGC14139
55	115166	AA258409	AF095727	Hs.287832	myelin protein zero-like 1
	115167	AA258421	AA749209	Hs.43728	hypothetical protein
	436719	AA262077	Y11192	Hs.5299	aldehyde dehydrogenase 5 family, member
	115239	AA278650	BE251328	Hs.73291	hypothetical protein FLJ10881
	115243	AA278766	AA806800	Hs.116665	KIAA1842 protein
60	428419	AA280791	U49436		KIAA1858 protein
	115322	AA280819	L08895	Hs.78995	MADS box transcription enhancer factor 2
	413303	AA280828	AW836130	Hs.75277	hypothetical protein FLJ13910
	115372	AA282195	AW014385	Hs.88678	ESTs, Weakly similar to Unknown (H.sapie
	409962	AA283127	U82671	Hs.57698	Target CAT
65	130269	AA284694	F05422	Hs.168352	nucleoporin-like protein 1
	456570	AA291137	AA286914	Hs.183299	ESTs
	332675	AA291708	BE439944		ESTs
	407864	AA293495	AF069291	Hs.40639	chromosome 8 open reading frame 1
70	115536	AA347193	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act
	408789	AA398474	AA059412	Hs.47986	hypothetical protein MGC10940
	115575	AA398512	AA393254	Hs.43619	ESTs
	115601	AA400277	AA148894	Hs.48849	ESTs, Weakly similar to ALU4_HUMAN ALU S
	434428	AA400896	D14540	Hs.199160	myeloid/lymphoid or mixed-lineage leukem
	115683	AA410345	AF255910	Hs.54650	junctional adhesion molecule 2
75	115715	AA416733	BE395161	Hs.1390	proteasome (prosome, macropain) subunit,
	132952	AA425154	AI658580	Hs.61426	Homo sapiens mesenchymal stem cell prote

	115818	AA426573	AA486620	Hs.41135	endomucin-2
	409124	AA431418	AW292609	Hs.50727	N-acetylglucosaminidase, alpha- (Sanfil)
	115895	AA436182	AB033035	Hs.51965	KIAA1209 protein
	458073	AA437099	AA182669	Hs.45032	ESTs
5	115962	AA446585	AJ636361	Hs.178520	hypothetical protein MGC10702
	115967	AA446887	AJ745379	Hs.42911	ESTs
	115974	AA447224	BE513442	Hs.238944	hypothetical protein FLJ10631
	115985	AA447709	AA447709	Hs.268115	ESTs, Weakly similar to T08599 probable
	129254	AA453624	AA252468	Hs.1098	DKFZp434J1813 protein
10	448730	AA455044	BE384932	Hs.64313	ESTs, Weakly similar to AF257182.1 G-pro
	116095	AA456045	AA043429	Hs.62618	ESTs
	426856	AA460454	R19768	Hs.172788	ALEX3 protein
	116210	AA476494	BE622782	Hs.172788	ALEX3 protein
	116213	AA476738	AA292105	Hs.326740	hypothetical protein MGC10947
15	432645	AA481422	D14041	Hs.347340	H-2K binding factor-2
	116265	AA482595	BE297412	Hs.55189	hypothetical protein
	129334	AA485084	AW157022	Hs.343551	hypothetical protein FLJ22584
	116274	AA485431	AJ129767	Hs.182874	guanine nucleotide binding protein (G pr
	426002	AA486638	BE514376	Hs.165998	PAI-1 mRNA-binding protein
20	116331	AA491000	N41300	Hs.71816	Homo sapiens mRNA; cDNA DKFZp586N1720 (f
	116333	AA491250	AF155827	Hs.203983	hypothetical protein FLJ10339
	132994	AA505133	AA112748	Hs.279905	clone HQ0310 PRO0310p1
	418538	AA598447	BE244323	Hs.85951	exportin, tRNA (nuclear export receptor
	116391	AA599243	T86558	Hs.75113	general transcription factor IIIA
25	116394	AA599574	NM_006033	Hs.65370	lipase, endothelial
	134531	AA600153	AJ742845	Hs.110713	DEK oncogene (DNA binding)
	116417	AA609309	AW499664		Human clone Z3826 mRNA sequence
	116429	AA609710	AF191018	Hs.279923	putative nucleotide binding protein, est
	116439	AA610068	AA251594	Hs.43913	PIBF1 gene product
30	116459	AA621399	R80137	Hs.302738	Homo sapiens cDNA: FLJ21425 fls, clone C
	427505	AA621752	AA361562	Hs.178761	26S proteasome-associated pad1 homolog
	409633	C21523	AW449822	Hs.55200	ESTs
	116541	D12160	D12160	Hs.249212	polymerase (RNA) III (DNA directed) (155
	132557	D19708	AA114926	Hs.169531	ESTs
35	414964	D25801	AA337548	Hs.333402	hypothetical protein MGC12760
	116571	D45652	D45652	Hs.211604	gbtHUMGS02848 Human adult lung 3' drect
	451522	D60208	BE565817	Hs.264698	hypothetical protein FLJ21657
	421919	D80504	AJ224901	Hs.109526	zinc finger protein 198
40	116643	F03010	AI367044	Hs.153638	myeloid/lymphoid or mixed-lineage leukemia
	116661	F04247	R61504		gbvht16a03.s1 Soares Infant brain 1N1B H
	116715	F10966	AL117440	Hs.170263	tumor protein p53-binding protein, 1
	116729	F13700	BE549407	Hs.115823	ribonuclease P, 40kD subunit
	318709	H05063	R52576	Hs.285280	Homo sapiens cDNA: FLJ22096 fls, clone H
45	418999	H16758	NM_000121	Hs.89548	erythropoietin receptor
	116773	H17315	AJ823410	Hs.343581	karyopherin alpha 1 (importin alpha 5)
	116780	H22566	H22566	Hs.63931	ESTs
	453884	H48459	AA355925	Hs.36232	KIAA0186 gene product
	116819	H53073	H53073	Hs.93698	EST
50	427278	H56559	AL031428	Hs.174174	KIAA0601 protein
	407833	H57957	AW956832	Hs.66666	ESTs, Weakly similar to S19560 proline-r
	116844	H64938	H64938	Hs.337434	ESTs, Weakly similar to A46010 X-linked
	116845	H64973	AA649530	Hs.348148	gbns44ID5.s1 NCL CGAP_Alv1 Homo sapiens
	116892	H69535	AI573283	Hs.38458	ESTs
55	116925	H73110	H73110	Hs.260603	ESTs, Moderately similar to A47582 B-cell
	116981	H81783	N29218	Hs.40280	ESTs
	453133	H86259	AC005757	Hs.31809	hypothetical protein
	117031	H88353	H88353	Hs.347265	gbvhw21a02.s1 Morton Fetal Cochlea Homo
	117034	H88639	U72209		YY1-associated factor 2
60	431129	H88675	AL137751	Hs.263671	Homo sapiens mRNA; cDNA DKFZp434J0812 (f
	417861	H93708	AA334551		sperm specific antigen 2
	117280	N22107	M18217	Hs.172129	Homo sapiens cDNA: FLJ21409 fls, clone C
	117344	N24046	R19085	Hs.210706	Homo sapiens cDNA FLJ13182 fls, clone NT
	117422	N27028	AI355562	Hs.43880	ESTs, Weakly similar to A46010 X-linked
65	117475	N30205	N30205	Hs.93740	ESTs, Weakly similar to I38022 hypotheti
	117487	N30621	N30621	Hs.44203	ESTs
	117937	N33258	AF044209	Hs.144904	nuclear receptor co-repressor 1
	130207	N33258	AF044209	Hs.144904	nuclear receptor co-repressor 1
	117549	N33390	N33390	Hs.44483	EST
70	117683	N40180	N40180		gbvry44d02.s1 Soares_multiple_sclerosis_
	117710	N45198	N45198	Hs.47248	ESTs, Highly similar to similar to Cdc14
	117791	N48325	N48325	Hs.93956	EST
	117822	N48913	AA706282	Hs.93963	ESTs
	422544	N49394	AB018259	Hs.118140	KIAA0716 gene product
75	117895	N50656	AW450348	Hs.93996	ESTs, Highly similar to SORL1_HUMAN SORTI
	452259	N50721	AA317439	Hs.28707	signal sequence receptor, gamma (translo
	133057	N53143	AA465131	Hs.64001	Homo sapiens clone 25218 mRNA sequence

	118103	N55326	AA401733	Hs.184134	ESTs
	118111	N55493	N55493		gbjv50c02.s1 Soares fetal liver spleen
	118129	N57493	N57493		gbjy54c08.s1 Soares_multiple_sclerosis_
5	118278	N62955	N62955	Hs.316433	Homo sapiens cDNA FLJ11375 fis, clone HE
	118329	N63520	N63520		gbjy6201.s1 Soares_multiple_sclerosis_
	118336	N63604	BE327311	Hs.47166	HT021
	417098	N64168	AB017365	Hs.173859	itizzed (Drosophila) homolog 7
	118363	N64168	AI183838	Hs.48938	hypothetical protein FLJ21802
	118364	N64191	N46114	Hs.29169	hypothetical protein FLJ22623
10	118475	N66845	N66845		gbza46c11.s1 Soares fetal liver spleen
	118491	N67135	AV647908	Hs.90424	Homo sapiens cDNA: FLJ23285 fis, clone H
	118500	N67295	W32889	Hs.154329	ESTs
	118584	N68963	AW136928		gbJIH-BI1-actp-d-08-0-UI.s1 NCL CGAP_Su
15	456647	N69331	AJ252640	Hs.110364	peptidylprolyl isomerase C (cyclophilin
	118661	N70777	AL137554	Hs.49927	protein kinase NYD-SP15
	118684	N71364	N71313	Hs.163986	Homo sapiens cDNA: FLJ22765 fis, clone K
	118689	N71545	AW390601	Hs.184544	Homo sapiens, clone IMAGE:3355383, mRNA,
	118690	N71571	N71571	Hs.269142	ESTs
	118766	N74456	N74456	Hs.50499	EST
20	118793	N75594	N75594	Hs.285921	ESTs, Moderately similar to T47135 hypot
	118817	N79035	AI668658	Hs.50797	ESTs
	118844	N80279	AL035364	Hs.50891	hypothetical protein
	118919	N91797	AW452696	Hs.130760	myosin phosphatase, target subunit 2
25	129558	N92454	AW580922	Hs.180446	karyopherin (importin) beta 1
	407604	N94581	AW191962	Hs.288061	collagen, type VIII, alpha 2
	118966	N94746	N94746	Hs.274248	hypothetical protein FLJ20758
	119021	N98238	N98238	Hs.55185	ESTs
	119039	R02384	AI160670	Hs.252097	pregnancy specific beta-1-glycoprotein 6
30	119063	R16833	R16833	Hs.53106	ESTs, Moderately similar to ALU1_HUMAN A
	332622	R41828	R10674		CSR1 protein
	119111	R43203	T02865	Hs.328321	EST
	415115	R46395	AA214228	Hs.127751	hypothetical protein
	119146	R58863	R58863	Hs.91815	ESTs
35	449224	R78248	AW995911	Hs.299883	hypothetical protein FLJ23399
	119239	T11483	T11483		gbCHR90049 Chromosome 9 exon Homo sapie
	119281	T16896	AI692322	Hs.65373	ESTs, Weakly similar to T02345 hypotheti
	119298	T23820	NM_001241	Hs.155478	cyclin T2
	126502	T30222	T10077	Hs.13453	hypothetical protein FLJ14753
40	419983	W15275	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFp586E1624 (f
	119558	W38194	W38194		Empirically selected from AFX single pr
	426641	W42414	AW081883	Hs.211578	Homo sapiens cDNA: FLJ23037 fis, clone L
	419445	W49632	AA884471	Hs.90449	Human clone Z3908 mRNA sequence
	119650	W57613	R82342	Hs.79856	ESTs, Weakly similar to S6557 alpha-1C-
45	119654	W57759	W57759		gbzdz20g11.s1 Soares_fetal_heart_NbHH19W
	119683	W61118	W65379	Hs.57835	ESTs
	119694	W65344	AA041350	Hs.57847	ESTs, Moderately similar to ICE4_HUMAN C
	119718	W69216	W69216	Hs.92848	ESTs
	410365	W69379	AJ287518		Homo sapiens mRNA; cDNA DKFp586D0923 (f
50	119938	W86728	AW014862	Hs.58885	ESTs
	120128	Z38499	BE379320	Hs.91448	MKP-1 like protein tyrosine phosphatase
	120130	Z38630	AA045767	Hs.5300	bladder cancer associated protein
	120148	Z39494	F02805	Hs.65765	ESTs
	120165	Z39623	Z39623	Hs.65783	ESTs
55	451979	Z40071	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	120183	Z40174	AW082866	Hs.65882	ESTs
	120184	Z40182	Z40182	Hs.65885	EST
	120211	Z40904	Z40904	Hs.66012	EST
	120245	AA166965	AW999815	Hs.111045	ESTs
60	120247	AA167500	AA167500	Hs.103939	EST
	120254	AA169599	W90403	Hs.111054	ESTs
	120259	AA171724	AW014786	Hs.192742	hypothetical protein FLJ12785
	120260	AA171739	AK000081	Hs.101590	hypothetical protein
	120275	AA177105	AA177105	Hs.78457	solute carrier family 25 (mitochondrial
65	120284	AA182626	AA179656		gbzps4e11.s1 Stratagene NT2 neuronal pr
	417735	AA186324	AA188175	Hs.82506	KIAA1254 protein
	422137	AA192099	AJ236885		zinc finger protein 148 (pH2-52)
	120302	AA192173	AA837098	Hs.269933	ESTs
	120303	AA192415	AJ216292	Hs.98184	ESTs
	120305	AA192553	AW295096	Hs.101337	uncoupling protein 3 (mitochondrial, pro
70	120319	AA194851	T57776	Hs.181094	ESTs
	408729	AA195520	AA195764	Hs.72639	ESTs
	120326	AA196300	AA196300	Hs.21145	hypothetical protein RG083M05.2
	133145	AA196549	H94227	Hs.6592	Homo sapiens, clone IMAGE:2961368, mRNA,
	120327	AA196721	AK000292	Hs.130732	hypothetical protein FLJ20285
75	120328	AA196979	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapi
	120340	AA206828	AA206828		gbzps80b08.s1 Stratagene hNT neuron (937

	417122	AA207123	AI906291	Hs.81234	Immunoglobulin superfamily, member 3
	131622	AA214539	AI380040	Hs.239489	TIA1 cytotoxic granule-associated RNA-bi
	421787	AA226914	AA227068	Hs.108301	nuclear receptor subfamily 2, group C, m
	120375	AA227260	AF028706	Hs.111227	Zic family member 3 (odd-paired Drosophi
5	120376	AA227469	AA227469		gb:zr18a07.s1 Stratagene NT2 neuronal pr
	120380	AA233122	AA837093	Hs.111460	calcium/calmodulin-dependent protein kin
	410804	AA233334	U84820	Hs.96521	Machado-Joseph disease (spinocerebellar
	434223	AA233347	AI825842	Hs.3778	zinc finger protein 216
	312771	AA233714	AA018515	Hs.264482	Homo sapiens mRNA; cDNA DKFZp761A0411 (f
10	120396	AA233796	AA134008	Hs.79306	eukaryotic translation initiation factor
	120409	AA235050	AA235050		gb:zs38e04.s1 Soares_NHMPu_S1 Homo sapi
	120414	AA235704	AW137156	Hs.181202	hypothetical protein FLJ110038
	120420	AA236031	AI128114	Hs.112885	spinal cord-derived growth factor-B
	120422	AA236352	AL133097	Hs.301717	hypothetical protein DKFZp434N1928
15	419326	AA236390	W94915	Hs.42419	ESTs
	120423	AA236453	AA236453	Hs.18978	Homo sapiens cDNA: FLJ22822 fis, clone K
	120435	AA243370	AA243370	Hs.96450	EST
	120453	AA250947	AA250947	Hs.170263	tumor protein p53-binding protein, 1
20	120455	AA251083	AA251720	Hs.104347	ESTs, Weakly similar to ALUC_HUMAN IIII
	120456	AA251113	AA488750	Hs.88414	BTB and CNC homology 1, basic leucine zi
	120473	AA251973	AA251973	Hs.269388	ESTs
	128922	AA252023	AI244901	Hs.9589	ubiquitin 1
	120477	AA252414	AA252414	Hs.43141	DKFZP727C091 protein
	120479	AA252650	AF006689	Hs.110299	mitogen-activated protein kinase kinase
25	120488	AA255523	AW952916	Hs.63510	KIAA0141 gene product
	120510	AA258128	AI796395	Hs.111377	ESTs
	120527	AA262105	AA262105	Hs.4094	Homo sapiens cDNA FLJ14208 fis, clone NT
	120528	AA262107	AI923511	Hs.104413	ESTs
	120529	AA262235	AI434823	Hs.104415	ESTs
30	120541	AA278298	W07318	Hs.240	M-phase phosphoprotein 1
	120544	AA278721	BE548277	Hs.103104	ESTs
	120562	AA280036	BE244580	Hs.342307	hypothetical protein FLJ10330
	120569	AA280648	AA807544	Hs.24970	ESTs, Weakly similar to B34323 GTP-bindl
35	120571	AA280738	AB037744	Hs.34892	KIAA1323 protein
	120572	AA280794	H39599	Hs.294008	ESTs
	129434	AA280837	AW967495	Hs.188644	ESTs
	130529	AA280886	AA178953	Hs.309648	gb:zp39e03.s1 Stratagene muscle 937209 H
	120575	AA280934	AW978022	Hs.238911	hypothetical protein DKFZp762E1511; KIAA
40	408339	AA281535	AB020686	Hs.54037	ectonucleotide pyrophosphatase/phosphodi
	120591	AA281797	AF076847	Hs.191356	general transcription factor IIH, polype
	120593	AA282047	AA748355	Hs.193522	ESTs
	430275	AA283002	Z11773	Hs.237786	zinc finger protein 187
	440303	AA283709	AA306166	Hs.7145	calpain 7
45	120609	AA283902	AW978721	Hs.268076	ESTs, Weakly similar to A46010 X-linked
	409702	AA284108	AI752244		eukaryotic translation elongation factor
	456870	AA284109	AI241084	Hs.154353	nonselective sodium potassium/proton exc
	132614	AA284371	AA284371	Hs.118064	similar to rat nuclear ubiquitous caselin
50	458750	AA284744	AA115496	Hs.336898	Homo sapiens, Similar to RIKEN cDNA 1810
	135376	AA284784	BE617856	Hs.99755	mitochondrial ribosome recycling factor
	120621	AA284840	AW961294	Hs.143818	hypothetical protein FLJ23459
	452279	AA286844	AA286844	Hs.61260	hypothetical protein FLJ13164
	332484	AA287032	AW172431	Hs.13012	ESTs
	120644	AA287038	AI869129	Hs.96616	ESTs
	120660	AA287546	AA286785	Hs.99677	ESTs
55	135370	AA287553	BE622187	Hs.99670	ESTs, Weakly similar to I38022 hypotheti
	120661	AA287556	AA287556	Hs.263412	ESTs, Weakly similar to ALUB_HUMAN IIII
	429828	AA287564	AB019494	Hs.225767	IDN3 protein
	452291	AA291015	AF015592	Hs.28853	CDC7 (cell division cycle 7, S. cerevisi
60	120699	AA291716	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos
	100690	AA291749	AA383256	Hs.1657	estrogen receptor 1
	120726	AA293656	AA293655	Hs.21198	ESTs
	120737	AA302430	AL049176	Hs.82223	chordin-like
	120745	AA302809	AA302809		gb:EST10426 Adipose tissue, white I Homo
65	443574	AA302820	U83993	Hs.321709	purinergic receptor P2X, ligand-gated io
	120750	AA310499	AI191410	Hs.96693	ESTs, Moderately similar to 2109260A B c
	120761	AA321890	AA321890		branched chain keto acid dehydrogenase E
	120768	AA340589	AA340589	Hs.104560	EST
	120769	AA340622	AI769467	Hs.9475	ESTs
70	135232	AA342457	AL038812	Hs.96800	ESTs, Moderately similar to ALU7_HUMAN A
	120793	AA342864	AA342864	Hs.96812	ESTs
	120796	AA342973	AI247356	Hs.96820	ESTs
	120809	AA346495	AA346495		gb:EST52657 Fetal heart II Homo sapiens
	332633	AA347573	AL120071	Hs.48998	fibronectin leucine rich transmembrane p
	120825	AA347614	AI280215	Hs.96885	ESTs
75	120827	AA347717	AA382525	Hs.132967	Human EST clone 122887 mariner transpos
	120839	AA348913	AA348913		gb:EST55442 Infant adrenal gland II Homo

	120850	AA349647	AA349647	Hs.96927	Homo sapiens cDNA FLJ12573 fis, clone NT
	120852	AA349773	AA349773	Hs.191564	ESTs
	128852	AA350541	R40622	Hs.106601	ESTs
	135240	AA357159	AA357159	Hs.96988	EST
5	120870	AA357172	AA357172	Hs.292581	ESTs, Moderately similar to ALU1_HUMAN A
	120894	AA370132	AA370132	Hs.97063	ESTs
	435737	AA370472	AF229839	Hs.173202	Kappa-B-interacting Ras-like protein 1
	120897	AA370867	AA370867	Hs.97079	ESTs, Moderately similar to AF174605 1 F
	120915	AA377296	AL135558	Hs.97104	ESTs
10	120935	AA383902	AL048409	Hs.97177	ESTs, Weakly similar to ALU1_HUMAN ALU S
	120936	AA385934	AA385934	Hs.97184	EST, Highly similar to (define not avai
	120937	AA386255	AA386255	Hs.97188	EST
	120938	AA386260	AA386260	Hs.104632	EST
	417632	AA386268	R20855	Hs.5422	glycoprotein M6B
15	120960	AA398014	AA398014	Hs.104684	EST
	120985	AA398222	AI219896	Hs.97592	ESTs
	120988	AA398235	AA398235	Hs.97631	ESTs
	121008	AA398348	AA398348	Hs.130546	Human DNA sequence from clone RP11-251J8
	121029	AA398482	AA398482	Hs.97641	EST
20	121032	AA398504	AA393037	Hs.161798	ESTs
	121033	AA398505	AA398505	Hs.97360	ESTs
	121034	AA398507	AL389951	Hs.271623	nucleoporin 50kD
	121035	AA398523	AA398523	Hs.210579	ESTs
	121058	AA398625	AA398625	Hs.97391	ESTs
25	121060	AA398632	AA398632	Hs.97395	ESTs
	121061	AA398633	AA393288	Hs.97396	ESTs
	121091	AA398894	AA398894	Hs.97657	ESTs, Moderately similar to ALU8_HUMAN A
	121092	AA398895	AA398895	Hs.97658	EST
	121094	AA398900	AA402505		gb:zd62h10.r1 Soares_testis_NHT Homo sap
30	121096	AA398904	AA398904	Hs.332690	ESTs
	121115	AA399122	AA398187	Hs.104682	ESTs, Weakly similar to mitochondrial cl
	121121	AA399371	AA399371	Hs.189095	similar to SALL1 (sal (Drosophila)-like
	121122	AA399373	AI126713	Hs.192233	ESTs, Highly similar to T00337 hypotethi
	121125	AA399441	AL042981	Hs.251278	KJAA1201 protein
35	121151	AA399636	AA399636	Hs.143629	ESTs
	121153	AA399640	AA399640	Hs.97694	ESTs
	121163	AA399680	AI676062	Hs.111902	ESTs
	121176	AA400080	AL121523	Hs.97774	ESTs
	121192	AA400262	AA400262	Hs.190093	ESTs
40	121223	AA400725	AI002110	Hs.97169	ESTs, Weakly similar to dJ667H12.2.1 [H]
	121227	AA400748	AA400748	Hs.97823	Homo sapiens mRNA; cDNA DKFZp434D024 (tr
	121231	AA400780	AA814948	Hs.96343	ESTs, Weakly similar to ALUC_HUMAN IIII
	121278	AA401631	AA037121	Hs.98518	Homo sapiens cDNA FLJ11490 fis, clone HE
	121279	AA401688	AA292873	Hs.177896	ESTs
45	121282	AA401695	AA401695	Hs.97334	ESTs
	121299	AA402227	AA402227	Hs.22826	tropomodulin 3 (ubiquitous)
	121301	AA402329	NM_006202	Hs.89901	phosphodiesterase 4A, cAMP-specific (dun
	121302	AA402398	AA402587	Hs.325520	LAT1-3TM protein
	121304	AA402449	AA293863	Hs.97316	EST
50	121305	AA402468	AA402468	Hs.291557	ESTs
	134721	AA403268	AK000112	Hs.89306	hypothetical protein FLJ20105
	121323	AA403314	AA291411	Hs.97247	ESTs
	121324	AA404229	AA404229	Hs.97842	EST
	444422	AA404260	AI768623	Hs.108264	ESTs
55	131074	AA404271	U16125	Hs.181581	glutamate receptor, ionotropic, kainate
	121344	AA405026	AA405026	Hs.193754	ESTs
	121348	AA405182	AA405182	Hs.97973	ESTs
	121350	AA405237	AA405237		gb:zd06e10.s1 NCLCGAP_GCB1 Homo sapiens
60	121400	AA406061	AA406061	Hs.98001	EST
	121402	AA406063	AA406063	Hs.98003	ESTs
	121403	AA406070	AA406070	Hs.98004	EST
	121408	AA406137	AA406137	Hs.98019	EST
	121431	AA406335	AA035279	Hs.176731	ESTs
	121471	AA411804	AA411804	Hs.261575	ESTs
65	121474	AA411833	AA402335	Hs.188760	ESTs, Highly similar to Trad [H.sapiens]
	121526	AA412219	AW665325	Hs.98120	ESTs
	121530	AA412259	AA778658	Hs.98122	ESTs
	121558	AA412497	AA412497		gb:zd95g12.s1 Soares_testis_NHT Homo sap
	121559	AA412498	AI192044	Hs.104778	ESTs
70	121584	AA416586	AI024471	Hs.98232	ESTs
	121609	AA416867	AA416867	Hs.98185	EST
	121612	AA416874	AA416874	Hs.98168	ESTs
	121737	AA421133	AA421133	Hs.104671	erythrocyte transmembrane protein
	121740	AA421138	AA421138	Hs.143835	EST
75	436032	AA422079	AA150797	Hs.109276	latexin protein
	121784	AA423837	T90789	Hs.94308	RAB35, member RAS oncogene family

	121802	AA424328	AI251870	Hs.188898	ESTs
	121803	AA424339	AI338371	Hs.157173	ESTs
	135286	AA424469	AW023482	Hs.97849	ESTs
	332778	AA424469	AW023482	Hs.97849	ESTs
5	121806	AA424502	AA424313	Hs.98402	ESTs
	129517	AA425004	AW972853	Hs.112237	ESTs
	121845	AA425734	AI732692	Hs.165066	ESTs, Moderately similar to ALU2_HUMAN A
	121853	AA425887	AA425887	Hs.98502	hypothetical protein FLJ14303
10	121891	AA426456	AA426456	Hs.98469	ESTs
	121895	AA427396	AA427396		gb:z33a02.s1 Soares ovary tumor NbHOT H
	121899	AA427555	R55341	Hs.50421	KIAA0203 gene product
	121917	AA428218	AA406397	Hs.139425	ESTs
	121918	AA428242	BE274689	Hs.184175	chromosome 2 open reading frame 3
	121919	AA428281	AA428281	Hs.98560	EST
15	121941	AA428865	AA428865	Hs.98563	ESTs
	121942	AA428994	AW452701	Hs.293237	ESTs
	121970	AA429668	AA429668	Hs.98617	EST
	121993	AA430181	AW297880	Hs.98661	ESTs
	418706	AA430184	U73524	Hs.87465	ATP/GTP-binding protein
20	122022	AA431293	AA431293	Hs.98716	ESTs, Moderately similar to T42650 hypot
	122050	AA431478	AI453076		ELAV (embryonic lethal, abnormal vision,
	122051	AA431492	AA431492	Hs.98742	EST
	122055	AA431732	AA431732	Hs.98747	EST
25	122105	AA432278	AW241685	Hs.98699	ESTs
	122125	AA434411	AK000492	Hs.98806	hypothetical protein
	135235	AA435512	AW298244	Hs.266195	ESTs
	122162	AA435698	AA626233	Hs.79946	cytochrome P450, subfamily XIX (aromatiz
	422072	AA435711	AB018255	Hs.111138	KIAA0712 gene product
30	415106	AA435815	U40763	Hs.77965	peptidyl-prolyl isomerase G (cyclophilin
	122186	AA435842	AA398811	Hs.104673	ESTs
	122235	AA436475	AA436475	Hs.112227	membrane-associated nucleic acid binding
	412970	AA436489	AB026436	Hs.177534	dual specificity phosphatase 10
	419288	AA442060	AA256106	Hs.87507	ESTs
	122310	AA442079	AW182803	Hs.98974	ESTs, Weakly similar to S65824 reverse t
35	122334	AA443151	BE465894	Hs.98365	ESTs, Weakly similar to LB4D_HUMAN NADP-
	122382	AA446133	AA446440	Hs.98643	ESTs
	122425	AA447145	AB007859	Hs.100855	KIAA0399 protein
	122431	AA447398	AA447398	Hs.99104	ESTs
40	122450	AA447643	AA447643	Hs.112095	hypothetical protein DKFZp434F1819
	426284	AA447742	AJ404468	Hs.284259	dymeh, axonemal, heavy polypeptide 9
	122477	AA448226	AA448226	Hs.324123	ESTs
	122500	AA448825	AA448825	Hs.99190	ESTs
	122522	AA449444	AA299607	Hs.98969	ESTs
45	122536	AA450087	AF060877	Hs.99236	regulator of G-protein signalling 20
	122538	AA450211	AA450211	Hs.99239	ESTs
	122540	AA450244	AA476741	Hs.98279	ESTs, Weakly similar to A43932 mucin 2 p
	122560	AA451213	AW392342	Hs.283077	centrosomal P4.1-associated protein; unc
	421919	AA452155	AJ224901	Hs.109526	zinc finger protein 198
50	122562	AA452156	AA452156		gb:z33a03.s1 Soares_totat_fetus_Nb2HF8_
	122585	AA453036	AI681654	Hs.170737	hypothetical protein FLJ23251
	122608	AA453526	AA453525	Hs.143077	ESTs
	122635	AA454085	AA454085		gb:z33a08.s1 Soares_totat_fetus_Nb2HF8_
	122636	AA454103	AW651706	Hs.99519	hypothetical protein FLJ14007
55	122653	AA454642	AW009166	Hs.99376	ESTs
	122660	AA454935	AI816827	Hs.180099	nuclear respiratory factor 1
	122703	AA456323	AA456323	Hs.269369	ESTs
	122724	AA457395	AA457395	Hs.99457	ESTs
	122749	AA458850	AA458850	Hs.293372	ESTs, Weakly similar to B34087 hypotheti
	122772	AA459662	AW117452	Hs.99489	ESTs
60	430242	AA459668	U66689	Hs.236642	3-hydroxyisobutyryl-Coenzyme A hydrolase
	429838	AA459679	AW904907	Hs.30732	hypothetical protein FLJ13409; KIAA1711
	122777	AA459702	AK001022	Hs.214397	hypothetical protein FLJ10160 similar to
	135362	AA460017	AA978128	Hs.99513	ESTs, Weakly similar to T17454 diaphanou
	122798	AA460324	AW366286	Hs.145696	splicing factor (CC1.3)
65	122837	AA461509	AA461509	Hs.293565	ESTs, Weakly similar to putative p150 [H
	122860	AA464414	AA464414		gb:z378g01.s1 Soares ovary tumor NbHOT H
	122861	AA464428	AA335721	Hs.213628	ESTs
	122910	AA470084	AA470084	Hs.98358	ESTs
70	132899	AA476606	AA476606	Hs.59666	SMAD in the antisense orientation
	122967	AA478521	AA806187	Hs.289101	glucose regulated protein, 58kD
	422845	AA478523	AA317841	Hs.7845	hypothetical protein MGC2752
	123009	AA479949	AA535244	Hs.78305	RAB2, member RAS oncogene family
	128917	AA481252	AI365215	Hs.206097	oncogene TC21
75	123081	AA485351	AI815486	Hs.243901	Homo sapiens cDNA FLJ20738 fis, clone HE
	123133	AA487264	AA487264	Hs.154974	Homo sapiens mRNA; cDNA DKFZp667N064 (fr
	123184	AA489072	BE247767	Hs.18166	KIAA0870 protein

	332467	AA489630	NM_014700Hs.119004	KIAA0685 gene product
	123233	AA490225	AW974175 Hs.151875	ESTs, Weakly similar to MAPB_HUMAN MICRO
	123234	AA490227	NM_001938Hs.16697	down-regulator of transcription 1, TBP-b
5	123236	AA490255	AW968504 Hs.123073	CDC2-related protein kinase 7
	123255	AA490890	AA830335 Hs.105273	ESTs
	430015	AA490916	AW768399 Hs.106357	ESTs
	448892	AA490925	AF084535 Hs.22464	epilepsy, progressive myoclonus type 2,
	123259	AA490955	AJ744152 Hs.283374	ESTs, Weakly similar to CA15_HUMAN COLLA
	123284	AA495812	AA488988 Hs.293796	ESTs
10	123286	AA495824	AA495824 Hs.188822	ESTs, Weakly similar to A46010 X-linked
	123315	AA496369	AA496369 Hs.109154	gbzv37d10.s1 Soares ovary tumor NbHOT H
	457397	AA504125	AW969025 Hs.109154	ESTs
	433049	AA521473	AL076668 Hs.334884	SEC10 (S. cerevisiae)-like 1
	123421	AA598440	AA598440 Hs.291154	EST, Weakly similar to I38022 hypothetical
15	123449	AA598899	AL049325 Hs.112493	Homo sapiens mRNA; cDNA DKFZp564D036 (fr
	426981	AA599244	AL044675 Hs.173081	KIAA0530 protein
	409986	AA599694	NM_014777Hs.57730	KIAA0133 gene product
	123497	AA600037	AA765256 Hs.135191	ESTs, Weakly similar to unnamed protein
	123604	AA609135	AA609135 Hs.293076	ESTs
20	123712	AA609684	AA609684 Hs.334437	Homo sapiens cDNA: FLJ21543 fs, clone C
	123731	AA609839	AA620423 Hs.112862	gbz6201.s1 Stratagene lung carcinoma
	123800	AA620423	AA620423 Hs.112862	EST
	123841	AA620747	AA620747 Hs.112896	ESTs
	123929	AA621364	AA621364 Hs.112981	ESTs
25	123978	C20653	T89832 Hs.170278	ESTs
	133184	D20085	AA001021 Hs.6685	thyroid hormone receptor interactor 8
	123235	D20749	Z83844 Hs.5790	hypothetical protein dJ37E16.5
	435147	D51285	AL133731 Hs.4774	Homo sapiens mRNA; cDNA DKFZp761C1712 (f
	128695	D59972	NM_003478Hs.101299	culin 5
30	124029	F04112	F04112 Hs.312553	gbHSC2.H062 normalized infant brain cDN
	124057	F13604	AA902384 Hs.73853	bone morphogenetic protein 2
	449316	H01662	AJ609045 Hs.321775	hypothetical protein DKFZp434D1428
	130973	H05135	AJ638418 Hs.1440	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep
	124106	H12245	H12245 Hs.101770	gbym17a12.r1 Soares infant brain 1N18 H
35	124136	H22842	H22842 Hs.101770	EST
	124165	H30894	H30039 Hs.107674	ESTs
	429627	H43442	NM_015340Hs.2450	leucyl-tRNA synthetase, mitochondrial
	124178	H45996	BE463721 Hs.97101	putative G protein-coupled receptor
40	129948	H69281	AI537162 Hs.263988	ESTs
	452114	H69485	N22687 Hs.8236	ESTs
	124+D826254		H69899 Hs.69899	gb:yu70c12.s1 Welzmann Olfactory Epithel
	129056	H70627	AJ769958 Hs.108336	ESTs, Weakly similar to ALUE_HUMAN IIII
	427580	H73260	AK001507 Hs.44143	Homo sapiens clone FLB6914 PRO1821 mRNA,
	426793	H77531	X89887 Hs.172350	HIR (histone cell cycle regulation defec
45	124274	H80552	H80552 Hs.102249	EST
	128078	H80737	AI351010 Hs.102287	lysosomal
	457658	H93412	AW952124 Hs.13094	presenilins associated rhomboid-like pro
	124315	H94892	NM_005402Hs.288757	v-ral simian leukemia viral oncogene hom
	437712	H95843	X04588 Hs.85844	neurotrophic tyrosine kinase, receptor,
50	124324	H96552	H96552 Hs.159472	Homo sapiens cDNA: FLJ22224 fs, clone H
	452933	H97146	AW391423 Hs.288555	Homo sapiens cDNA: FLJ22425 fs, clone H
	132231	H99131	AA662910 Hs.42635	hypothetical protein DKFZp434K2435
	421877	H99462	AW250380 Hs.109059	mitochondrial ribosomal protein L12
	443123	H99837	AA094538 Hs.272808	putative transcription regulation nuclea
55	132963	N22140	AA099693 Hs.34851	epsilon-tubulin
	420473	N22197	AL118782 Hs.300208	Sec23-interacting protein p125
	417381	N23756	AF164142 Hs.82042	solute carrier family 23 (nucleobase tra
	130365	N24134	W56119 Hs.155103	eukaryotic translation initiation factor
	456610	N24195	AF172066 Hs.106346	retinoic acid repressible protein
60	439311	N26739	BE270668 Hs.151945	mitochondrial ribosomal protein L43
	124383	N27098	N27098 Hs.102463	EST
	124387	N27637	N27637 Hs.109019	ESTs
	129341	N33090	AI193519 Hs.226396	hypothetical protein FLJ11126
	419793	N35967	AI364933 Hs.168913	serine/threonine kinase 24 (Ste20, yeast
65	124433	N39089	AA280319 Hs.288840	PRO1575 protein
	124441	N46441	AW450481 Hs.161333	ESTs
	132338	N48270	AA353868 Hs.182982	golgin-57
	436575	N48365	AI473114 Hs.113319	ESTs
	124466	N51316	R10084 Hs.113319	kinesin heavy chain member 2
70	408048	N51499	NM_007203Hs.42322	A kinase (PRKA) anchor protein 2
	124483	N53976	AI821780 Hs.179864	ESTs
	124484	N54157	H66118 Hs.285520	ESTs, Weakly similar to 2109260A B cell
	124485	N54300	AB040933 Hs.15420	KIAA1500 protein
	124494	N54831	N54831 Hs.271381	ESTs, Weakly similar to I38022 hypotheti
75	129200	N59849	N59849 Hs.13565	Sam68-like phosphotyrosine protein, T-ST
	124527	N62132	N79264 Hs.269104	ESTs

	124532	N62375	N62375	Hs.102731	EST
	133213	N63138	AA903424	Hs.6786	ESTs
	124539	N63172	D54120	Hs.146409	cell division cycle 42 (GTP-binding prot
5	129196	N63787	BE296313	Hs.265592	ESTs, Weakly similar to I38022 hypothetical
	124575	N68168	N68168		gbca11c01.s1 Soares fetal liver spleen
	124578	N68201	N68201		ESTs, Weakly similar to I38022 hypothetical
	124577	N68300	N68300	Hs.138485	gbca12g07.s1 Soares fetal liver spleen
	124578	N68321	N68321	Hs.231500	EST
10	124593	N69575	N69575	Hs.102768	ESTs
	128501	N75007	AL133572	Hs.199009	protein containing CXXC domain 2
	332434	N75542	AI680737	Hs.289068	Homo sapiens cDNA FLJ11918 fis, clone HE
	128473	N90066	T78277	Hs.100293	O-linked N-acetylglucosamine (GlcNAc) tr
	128639	N91246	AW562962	Hs.102897	CGI-47 protein
15	124652	N92751	W19407	Hs.3862	regulator of nonsense transcripts 2; DKF
	133137	N93214	AB002316	Hs.65746	KIAA0318 protein
	124671	N99148	AK001357	Hs.102951	Homo sapiens cDNA FLJ10495 fis, clone NT
	133054	R07876	AA464836	Hs.291079	ESTs, Weakly similar to T27173 hypothetical
	425268	R10865	J00077	Hs.155421	alpha-fetoprotein
20	124720	R11056	R05283		gbrye91c08.s1 Soares fetal liver spleen
	124722	R11488	T97733	Hs.185685	ESTs
	128944	R23930	AL137586	Hs.52763	anaphase-promoting complex subunit 7
	132865	R26589	AI248173	Hs.191460	hypothetical protein MGC12835
	426504	R37588	AW162919	Hs.170160	RAB2, member RAS oncogene family-like
25	438828	R37613	AL134275	Hs.6434	hypothetical protein DKFZp761F2014
	124757	R38398	H11368	Hs.141055	Homo sapiens clone 23758 mRNA sequence
	124762	R39179	AA553722	Hs.92098	ESTs, Moderately similar to A46010 X-in
	124773	R40923	R45154	Hs.338439	ESTs
	135266	R41179	R41179	Hs.97393	KIAA0328 protein
30	427961	R41294	AW293165	Hs.143134	ESTs
	414303	R42307	NM_004427	Hs.165263	early development regulator 2 (homolog o
	128540	R43189	AW297929	Hs.328317	EST
	124785	R43306	W38537	Hs.280740	hypothetical protein MGC3040
	124792	R44357	R44357	Hs.48712	hypothetical protein FLJ20736
35	124793	R44519	R44519		gbryg24h04.s1 Soares infant brain 1NIB H
	124799	R45088	R45088		gbryg38g04.s1 Soares infant brain 1NIB H
	124812	R47948	R47948	Hs.188732	ESTs
	124821	R51524	H87832	Hs.7388	kelch (Drosophila)-like 3
40	424123	R54950	AW966158	Hs.58582	Homo sapiens cDNA FLJ12789 fis, clone NT
	124835	R55241	R55241	Hs.101214	EST
	124845	R59585	R59585	Hs.101255	ESTs
	124847	R60044	W07701	Hs.304177	Homo sapiens clone FLB8503 PRO2286 mRNA,
	440630	R60872	BE561430	Hs.239388	Human DNA sequence from clone RP1-304B14
	124861	R66690	R67567	Hs.107110	ESTs
45	332503	R67266	NM_004455	Hs.150956	exostoses (multiple)-like 1
	124879	R73588	R73588	Hs.101533	ESTs
	124892	R79403	AI970003	Hs.23756	hypothetical protein similar to swine ac
	124906	R87647	H75964	Hs.107815	ESTs
	124922	R93622	R93622	Hs.12163	eukaryotic translation initiation factor
50	124940	R99599	AF068846	Hs.103804	heterogeneous nuclear ribonucleoprotein
	124941	R99612	AI766661	Hs.27774	ESTs, Highly similar to AF161349 1 HSPC0
	124943	T02888	AW963279	Hs.123373	ESTs, Weakly similar to ALU1_HUMAN ALU S
	124947	T03170	T03170	Hs.100165	ESTs
	124954	T10465	AW964237	Hs.67728	KIAA1548 protein
55	456862	T15418	U55184	Hs.154145	hypothetical protein FLJ11585
	410653	T15597	BE383768	Hs.65238	95 kDa retinoblastoma protein binding pr
	418133	T15652	R43504	Hs.6181	ESTs
	440014	T16896	AW960782	Hs.6856	ash2 (absent, small, or homeotic, Drosop
	131082	T26644	AI091121	Hs.246218	Homo sapiens cDNA: FLJ21781 fis, clone H
	124980	T40841	T40841	Hs.98681	ESTs
60	124984	T47566	BE313210	Hs.334798	eukaryotic translation elongation factor
	124991	T50116	T50116		gbryb77c10.s1 Stratagene ovary (937217)
	457222	T50145	NM_004477	Hs.203772	F5HD region gene 1
	125000	T58615	T58615	Hs.235887	ESTs
65	132932	T59940	AW118826	Hs.6093	Homo sapiens cDNA: FLJ22783 fis, clone K
	444484	T63595	AK002126	Hs.11260	hypothetical protein FLJ11264
	125008	T64891	T91251		gbryd60a10.s1 Soares fetal liver spleen
	125009	T64924	T64924	Hs.303046	ESTs
	445384	T64933	T79136	Hs.127243	Homo sapiens mRNA for KIAA1724 protein,
70	125017	T68875	T68875		gbryc30R05.s1 Stratagene liver (937224)
	125018	T69027	T69027	Hs.269481	sex comb on midleg homolog 1
	125020	T69924	T69981		gbryc19d03.r1 Stratagene lung (937210) H
	437871	T70353	AI084813	Hs.114088	ESTs
	134204	T79780	AI873257	Hs.7994	hypothetical protein FLJ20551
75	125050	T79951	AW970209	Hs.111805	ESTs
	125052	T80174	T85104	Hs.222779	ESTs, Moderately similar to similar to N
	125054	T80622	T80622	Hs.268601	ESTs, Weakly similar to envelope [H.sapi

	125063	T85352	T85352	gb-yd82d01.s1 Soares fetal liver spleen
	125064	T85373	T85373	gb-yd82d07.s1 Soares fetal liver spleen
	125068	T86284	T86284	gb-yd77b07.s1 Soares fetal liver spleen
5	416507	T89579	AL045364 Hs.79353	transcription factor Dp-1
	125080	T90360	T90360 Hs.268620	ESTs, Highly similar to ALU6_HUMAN ALU S
	125097	T94328	AW576389 Hs.335774	EST, Moderately similar to S65657 alpha-
	125104	T95590	T95590	gb-yd40a03.s1 Soares fetal liver spleen
	135107	T97257	T97257 Hs.94660	ESTs, Moderately similar to I38022 hypot
10	423122	T97599	AA845462 Hs.124024	deltex (Drosophila) homolog 1
	125118	T97620	R10606 Hs.269890	gb-yd35f11.s1 Soares fetal liver spleen
	125120	T97775	T97775 Hs.100717	EST
	134160	T98152	T98152 Hs.79432	fibrillin 2 (congenital contractural ara
	125136	W31479	AW962364 Hs.129051	ESTs
15	125144	W37999	AB037742 Hs.24336	KIAA1321 protein
	125150	W38240	W38240	Empirically selected from AFFX single pr
	450142	W40150	AW207469 Hs.24485	chondroitin sulfate proteoglycan 6 (bama
	131987	W45435	AW453069 Hs.3657	activity-dependent neuroprotective prote
	125178	W58202	W53127 Hs.31845	ESTs
20	125180	W58344	W58469 Hs.103120	ESTs
	125182	W58650	AA451755 Hs.263560	ESTs
	446888	W68736	AL030996 Hs.16411	hypothetical protein LOC57187
	125197	W69106	AF086270 Hs.278554	heterochromatin-like protein 1
	133497	W69111	BE617303 Hs.74266	hypothetical protein MGC4251
25	429922	W69399	Z97630 Hs.226117	H1 histone family, member 0
	129232	W69459	R98881 Hs.109655	sex comb on midleg (Drosophila)-like 1
	422166	W72424	W72424 Hs.112405	S100 calcium-binding protein A9 (calgran
	125209	W72724	W72724 Hs.103174	ESTs, Weakly similar to TSP2_HUMAN THROM
	125212	W72834	AA746225 Hs.103173	ESTs
30	456631	W73955	BE383436 Hs.108847	hypothetical protein MGC2749
	125223	W74701	AI916289 Hs.109057	ESTs, Weakly similar to ALU5_HUMAN ALU S
	125228	W76540	W74169 Hs.16492	DKFZP564G2022 protein
	125228	W79397	AA033982 Hs.110059	ESTs, Weakly similar to I38022 hypotheti
	132393	W85888	AL135094 Hs.47334	hypothetical protein FLJ14495
35	125238	W86038	N99713 Hs.109514	ESTs
	125247	W86881	AA694191 Hs.163914	ESTs
	129296	W87804	AI051967 Hs.110122	ESTs
	125263	W88942	AA088878	gb-zn45g10.r1 Stratagene HeLa cell s3 93
40	125268	W90022	W80022 Hs.186809	ESTs, Highly similar to LCT2_HUMAN LEUKO
	450862	W92272	U91543 Hs.25601	chromodomain helicase DNA binding protei
	452401	W92764	NM_007115 Hs.29352	tumor necrosis factor, alpha-induced pro
	428243	W93040	H05317 Hs.283549	ESTs
	125277	W93227	W93227 Hs.103245	EST
	125278	W93523	AI218439 Hs.129998	enhancer of polycomb 1
45	125280	W93659	AI123705 Hs.106932	ESTs
	448205	W94003	W93949 Hs.33245	ESTs
	131844	W94401	AI419294 Hs.324342	ESTs
	125284	W94688	NM_002666 Hs.103253	perlepin
	417111	W94787	AW016321 Hs.82306	desitin (actin depolymerizing factor)
50	445424	Z38294	AB028945 Hs.12696	cortactin SH3 domain-binding protein
	125289	Z38311	T34530 Hs.4210	Homo sapiens cDNA FLJ13069 fis, clone NT
	446313	Z38465	H06245 Hs.106801	ESTs, Weakly similar to PC4259 ferritin
	431342	Z38525	AW971018 Hs.21659	ESTs
	433227	Z38538	AB040923 Hs.106808	kelch (Drosophila)-like 1
55	428306	Z38551	AB037715 Hs.183639	hypothetical protein FLJ10210
	424624	Z38783	AB032947 Hs.151301	Ca2+-dependent activator protein for secr
	125295	Z39113	AB022317 Hs.25887	sema domain, immunoglobulin domain (lg),
	125298	Z39255	AW972542 Hs.289008	Homo sapiens cDNA: FLJ21814 fis, clone H
	125300	Z39591	Z39591 Hs.101376	EST
60	446378	Z39783	BE622770 Hs.264915	Homo sapiens cDNA FLJ12908 fis, clone NT
	444582	Z39920	R55344 Hs.22142	cytochrome b5 reductase b5R.2
	130882	Z40168	AA497044 Hs.20887	hypothetical protein FLJ10392
	128888	Z40388	AI760853 Hs.241558	ariadne (Drosophila) homolog 2
	125310	Z40646	R59161 Hs.124953	ESTs
65	125315	Z41697	R38110 Hs.106296	ESTs
	125317	Z99349	Z99348 Hs.112461	ESTs, Weakly similar to I38022 hypotheti
	135096	Z99394	AA081258	zinc finger protein 36 (KOX 18)

TABLE 3A

Table 3A shows the accession numbers for those pkeys lacking unigenelD's for Table 3. The pkeys in Table 7 lacking unigenelD's are represented within Tables 1-6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustalng and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
108469	116761_1	AA079487 AA128547 AA128291 AA079587 AA079600
124106	125446_1	H12245 AA094769 R14578
108501	13684_12	AA083258
108562	36375_1	AA100798 AF020589 AA074629 AA075946 AA100849 AA085347 AA126309 AA079311 AA079323 AA085274
101300	4669_1	BE535511 M62098 AA306787 AW891766 AA348998 AA338869 AA344013 AW956561 AW389343 AW403607 L40391
		AW408435 AA121738 A1568978 H13317 R20373 AW948724 AW948744 AA335023 AA336722 AA448690 C21404
		AW884390 AA345454 AA303292 AA174174 BE092280 T90614 AA035104 R76028 AA126924 AA741086 AW022056
		AW118940 AA121666 A1832408 AA683475 A1140901 A1623576 AW519064 AW474125 A1953923 A1735349 AW150109
		A1436154 AW118130 AW270782 A1804073 N27434 AA876543 AA937815 A1051166 AA505378 A1041975 A1335355
		A1089540 AA662243 A1127912 A1925604 A1250880 A1366874 A1564386 A1615196 A1683526 A1435885 A1160934 H79030
		A1801493 AA448691 A1673767 A1078042 A1804327 AA813438 AA680002 A1274492 T16177 A1287337 A1935050
		AA907805 AA911493 A1589411 A1371358 AW576236 A1078866 AW516168 AA346372 A1560185 AA471009 R75857
		AA296025 AA523155 AA853168 A1686593 A1658482 A1566601 AW072797 AA128047 AA035502 AW243274 AA992517
		R43760
132091	94851_1	AW954243 AA829930 AA412478 AA828434 AA814538 A1927418 A1192435 W52897 AA443666 AA031913 A1683306
		AA918481 A1183314 D83907 A1206832 AA876122 D83836 D83838 D82533 A1761290 A1191125 A1143749 AW771909
		A1241436 A1767267 W556507 AA847787 AA586692 T10502 A1247870 AA715017 AA643304 AA890233 AA811387
		AA897470 AA907729 A1070679 A1078010 AA452830 AW419160 A1783713 N80205 W56778 AA676899 A1888718 N69930
		A1338935 A1217580 AA639508 AA575836 BE046852 A1312651 A1038406 AA628649 AA643838 A1493761 AA032024
		W38849 AA340178 AA447052 AA452969 W19369 AA296364 H44229 W58767 C05751 C05835 A1741989 N98532
		AW102617 AA412583 A192246 W38496 AA355375 AA928571 C05275 AA352500 N93132
		U72209 NM_005748 A1655607 A1052758 AA385199 AW956794 H88679 AL135153 A1765644 AA384399 AW966458
		AA568443 AA804610 A1873513 H88639 Z25371 R63458 W44919
100752	33207_21	T81309 BE019033 R94181 BE019198 NM_000612 J03242 AW411299 BE300064 BE297544 R94182 AW630108 T53723
		D58853 H78073 H80594 BE299560 T48899 H70196 M17428 N77077 S77035 H58384 H61664 H78540 T84527 C17198
		H60255 H71880 R92644 W79050 X00910 M29645 R91055 M17863 M17862 T71815 BE299561 BE464561 X06260
		R94741 T54216 C18594 BE262015 X06161 AW409889 AA378400 BE263228 BE313278 R88116 BE313457 H43500
		T46617 BE313761 H77309 A1207601 X06159 H40413 X03425 T87663 R10627 X03562 M14118 W03982 R97520 H81229
		T83157 H83168 H48762 AA669898 BE263054 H47289 AA022807 R11555 H74260 R76968 R28338 H72534 H72464
		H62031 N72478 A15355 AW411300 R89113 R89135 H58454 T83281 R93476 H69645 H68015 T82228 H71089 T85121
		H59939 W65299 N78176 H53909 N72373 R21788 H04660 H59639 H61874 BE262219 T53614 N73335 N50464 W00943
		N77189 R89257 AA570502 R89432 R06366 AA553480 AA776271 AA551359 AA551050 H51670 AA601052 BE299081
		H68198 H52278 BE207832 N91192 H70332 X07868 X07868 H69464 H53782 H73710 R80435 AA553384 AW884176
		N53475 T71662 AW954036 AW954033 AA552931 H93206 AA430218 AA553476 A1918470 T54124 BE207982 BE300177
		N73994 AW882625 N39549 N53838 AA722389 H71878 H58909 H37849 H78435 T47933 R77174 R83614 AA411890
		H94199 AA663208 BE205778 AA490137 H70492 R98232 H37800 AA679294 H40341 H74238 H47290 H73231 T48618
		AA025428 A1039521 H92969 N53389 H80538 H72933 T90630 AA411891 N55000 H74225 AA340290 AW957061 T54316
		AA340437 H57125 H58908 H79027 H63450 N74623 R93425 H68714 H68758 H68396 H48763 H69256 H57320 H53831
		H53588 N68833 N52453 H56048 H69870 H78074 R69253 R83375 T53615 H94330 H58455 H90884 T47934 H74261
		R89258 R87997 R91058 R28339 R86760 H78235 R97521 H67692 H40358 AA022688 H52513 H59601 T88690 H65256
		H63397 W65397 AA553588 R19280 N52645 W73930 R06367 R21743 H72372 N73921 AW883539 AW882639 T40616
		H47084 R95723 AA634316 AA862781 H77310 R91389 H93111 R92767 T54512 R89341 H70333 H57817 H82941 H62032
		N52638 H58385 T91796 H51086 AA340292 T49918 H81230 R36121 N50411 T87664 N62436 N39340 AA665637
		AA340446 H93377 H92973 BE296290 BE269788 H61665 AA340444 N54605 AA454101 R10628 R94200 A1200549
		AA342640 BE298855 BE250229 T49916 H82008 N28278 AW880662 H71268 N76791 H47685 H55255 W05198
		AW889144 N76677 H71702 H68036 H71915 R91612 R87807 H68059 A1133328 A1247866 AA621443 AW881050
		AA700847 AA340413 AW878608 AW881181 AW878249 H71916 N54596 BE161581 AW878082 W04212 AW881040
		AW885492 AW880519 AA334887 AW878715 W06882 AW630222 AW885381 H70869 AW361778 H47501 AW889982
		H63668 AW864986 AW878713 AW878685 R36391 AW878694 AA368070 C03393 AW878695 AW878705 AW878665
		AW878742 AW878620 AW878823 AW878688 R29048 AW878690 AW878686 AW878810 AW878827 AW878733
		AW878659 AW878749 AW878681 AW883353 AW883277 AW883300 AW883565 AW883298 AW883143 AW883045
		AW883482 AW883352 AW883417 AW883357 AW883231 AW883474 AW883355 AW882620 AW882533 AW883754
		AW883139 AW882827 AW883641 AW883567 AW883481 AW882983 AW882982 AW882465 AW883419 AW882466
		AW883639 AW883230 AW882981 AW882534 AW882874 AW882619 AW883480 AW882826 AW882631 AW882835
		AW882830 AW883563 AW882456 AW627642
116417	5418_11	AW499664 AW500888 AL042095 AW576556 AW265424 A1521500 AA781333 AA761319 AW291137 AA649040 AA769094
		AA489864 AA635311 AW070509 AA425658 A1381489 AA609309 AA134476 W74704 A1923640 AW084888 H45700
		A1985564 AW629495 AW614573 A1859571 A1693486 AA913892 A1806164 AA909524 AW263513 A1356361 Z40708

			AJ332765 AJ392620 AA181060 AW118719 AW968804 AW263502 AW505314 AA036967 W74741 R51139 H19364 H45751 Z44962 AW370823 H25650 T54007 AA453000 AL045739 AA609684 AA758732 W73853 AA928112 W77887 AW889237 AA148524 AJ749182 AJ754442 AJ338392 AJ253102 AJ079403 AJ370541 AJ697341 H97538 AW188021 AJ927669 W72716 AJ051402 AJ188071 AJ335900 N21488 AW770478 W92522 AB91028 AJ913512 AJ144448 W73819 AA604358 N28900 W95221 AJ868132 H98465 AA148793 T91251 T84891 T85665 T69981 T69924 AA078476 T86284 T81933 R61504 F04247 T95590 AA703278 H62764 N68168 N69188 N90450 AA098878 W88942 AW980564 AA092457 T55890 D56120 T92525 AJ815967 BE182608 BE182595 AW080238 M90657 AA347236 AW961686 AW176446 AA304671 AW583735 T81714 AA316968 AJ446615 AA343532 AA083489 AA488005 W52095 W39480 N57402 D82638 W25540 W52847 D82729 D58990 BE819182 AA315188 AA308636 AA112474 W76162 AA088544 H52265 AA301631 H80982 AA113786 BE620987 AW651691 AA343799 BE613669 BE547180 BE546656 F11933 AA376800 AW239185 AA376086 BE544387 BE619041 AA452515 AA001806 AA190873 AA180483 AA159546 F00242 AJ940609 AJ940602 AJ189753 T87663 T66110 AW062896 AW062910 AW062902 AJ051622 AJ828930 AA102452 AJ685095 AJ819390 AA557597 AA383220 AJ804422 AJ633575 AW338147 AW603423 AW606800 AW750567 AW510672 AJ250777 AA083510 AW629109 AW513200 AA921353 AJ877934 AJ148698 AJ955858 AA173825 AA453027 AJ027865 AW375542 AA454099 AA733014 AJ591384 R79300 R80023 AA843108 AA626058 AA84898 AW375550 AA889018 AJ474275 AW205937 AJ052270 AW388117 AW388111 AA699452 AJ242230 N47478 H38178 AA366621 AA113196 AA130023 H39740 T61629 AJ85973 AW083671 AA179730 AA305757 AJ285455 N83956 AA216013 AA336155 AW999595 T97525 AA345349 T91762 AA771981 AJ285092 AJ581386 BE392486 BE385852 AA682601 AJ682884 AA345840 T85477 AA292949 AA932079 AA098791 D82607 T48574 AW752038 C06300 R20840 R20839 M30269 NM_002508 X82245 AJ078760 AW957003 D78945 M27445 AA650439 AL048816 AV660256 AV660347 AA333052 BE295257 T60999 AA383049 AW369677 Z26985 AW175704 AA343326 AW747957 AJ818389 W17308 W17302 H15591 AA371284 AA370412 W94966 BE384365 T28498 R80714 R16859 H21723 AW835154 D56097 D56381 W21232 AA180565 AW379755 AW067895 AW136928 AJ685655 BE218584 BE465078 N68963 AA975338 BE147199 N76377 BE273749 BE397561 BE387189 AL037858 AL037878 AJ963094 BE259216 AA011363 AL036189 BE562325 AA251169 BE617431 N96537 AA158093 AL047800 M34539 NML_000801 AA312140 D16971 AA158904 AA307114 AA312803 T09203 AW629686 AL048504 BE388578 AA220957 AA158364 BE267385 AA294971 C18055 BE241757 AA115056 AJ936769 BE378435 BE206971 AW674924 BE622060 AA604674 AA115273 AW402159 AA338608 BE568819 M80199 X55741 AA375111 AA376016 BE612671 AA805742 AW405588 N25850 AA4580 H06031 AW403549 BE336552 AA056726 BE543239 AA082517 AJ201645 AJ201642 AJ192622 N40104 AA370921 BE547569 AJ969602 AA302038 AJ197890 AW268354 AJ014938 W45448 AJ541395 AA037272 BE538826 AL039613 BE536130 AA299355 AW805147 AW574624 H53220 AJ47171 AA399303 AA007386 W35106 BE613277 R12739 R12738 AA304342 AA687802 BE408581 AJ498844 AV662092 AW904105 AA011375 BE315214 H93302 BE537893 N32289 AW855829 AJ291320 BE078322 AJ301395 AA303362 N32719 AA358328 AA357877 AJ952540 H58279 H02758 H02048 AW805233 R82224 AA10772 AA291352 BE171109 N69935 BE169248 AA361173 H44978 BE617887 D52560 AA084043 W03595 R67219 N36477 N42924 R67104 H44901 H79695 W21105 AA393988 W30899 AA316096 BE522896 W46872 AA442678 BE544893 BE540112 BE621873 AA338067 N55052 BE398154 BE621210 AA740760 C03739 C03206 BE396692 AA482370 AA031614 AA301575 AA304710 AA132153 AA029796 AA994960 H19567 AA442969 H49781 H46871 AA035395 AA056185 AJ49378 AA643080 AL135479 AA292329 AA654337 AA041228 AA454888 AA025039 W58331 AA625981 T94941 AA302448 H19900 AA218956 AA513790 AA563962 AA398078 W44441 AA293276 W47373 AA625879 W30688 AA043029 T64284 R78151 AA304340 AA485186 AA604939 R82470 AA421425 AW771456 AJ333329 AA304424 AA605236 AA938934 AA587673 AJ209162 AJ697301 AJ479995 AJ679814 AJ381950 AW189125 AJ955888 AJ986019 BE301019 AJ084792 AJ310211 AW189307 AJ022070 AW977204 AJ146825 AW190163 AW303281 AJ828345 BE046043 AW029257 AA482268 AJ246507 AJ420729 AW084932 AW439514 AJ890487 AW439692 AJ523896 AJ186612 AJ659953 AJ889773 AA687527 AW072694 AW262153 AW467371 AJ613269 AJ679238 D54404 AA158103 AW105527 AW149739 AW150361 AW268387 AW117708 AJ951682 AJ687440 AW674285 AJ678365 AJ587082 AA732095 AA019899 W45661 AA627300 BE613304 AA765891 AA612935 AJ814658 AW316916 R66594 AA514640 AA025040 AA031472 AW732076 AA029797 AJ244560 AJ128734 AW381720 AJ092360 AJ263283 AW613175 AJ890675 AJ720156 AW631348 AJ635106 AJ278045 AA303979 AA703505 W45449 AW078661 AJ292052 AW381707 AJ147854 AW381743 AA158905 AA303258 AA888144 AW195967 AA428708 AA989559 AA617731 H19882 BE543418 AA830386 AA421302 W58652 T94995 AJ869743 AJ679145 AW085971 N98425 AA765136 AJ347027 AJ356955 AA928038 AJ679717 AA458459 AA679281 AJ67973 AJ270041 AA765135 AA732793 AJ798447 AA668646 AA251008 AJ984538 AJ401737 AA056186 BE043308 AW662375 AJ302110 N50724 W96332 BE537047 N26983 AJ567172 AA765296 AW673237 N29784 AA534275 AA084044 AW067973 AW300766 T63398 W46823 R39790 AJ364185 AW298582 AA454814 AW069878 N67751 H05982 N23140 AJ362647 AJ302086 AJ767772 N25755 H53114 AA706133 T33511 AA429281 AA935294 AA987647 W02803 R66595 AJ680795 W23673 AW440794 AA722872 H49538 AW131042 AA531603 AA908665 AA040791 AA235312 W52205 N33444 R82180 H02759 H79696 AW088894 H56079 AA961143 AW067776 AW937345 AA016311 AW071227 AA017511 AJ753994 W47374 T84155 AA296092 AJ598626 AA558158 AA296088 AW794259 H01963 AA149267 AA485076 AA975856 H44938 AA035396 AJ955555 H46289 AA486161 AJ631222 AA359047 AW794253 AJ806962 AW243930 AA526145 AW878734 AA018464 AA132031 R67220 R79152 AA296093 H54300 AJ005160 BE242548 AW992803 AW878644 AW878666 T27742 R82471 AW517604 AW472738 AJ282904 R39791 AA486098 AW467891 AW960520 AA551736 AA056621 AW945197 R66373 AA564236 BE242202 AJ904376 AJ82590 H19484 R00890 AJ627677 AA302287 AJ869451 AJ734855 AJ708073 AJ832902 AA585184 AW204299 AA055565 D12417 D11975 T63543 AW664099 R54423 BE612712 T96340 T63985 AA598917 T40735 T64053 AA149284 AW272548 AA363445 AA042893 AW300697 BE261973 T53501 T53500 AW878729 AW878657 AW794391 AA069193 R01553 H44875 AA385408 AA533968 M3060 AL135600 W96331 AA017651 AA018849 AA017692 H85337 BE278690 AA731598 AA018512 AJ076813 AJ022644 R02585 X52220 AW296894 AA825671 AJ699321 AJ393601 AW592611 AJ146747 AA608921 AA158365 AW590007 AA354519 D20081 R02704 AW798339 M92422 AA094903 AA007676
--	--	--	--

WO 02/079492

PCT/US02/04915

133681	13893_1	AI352558 ZB2248 X78138 NM_003405 AU077248 AA223125 S80794 D78577 AH124697 AW403970 BE614089 BE296713 BE621334 L20422 X80536 D54224 D54950 X57345 N28228 AA127798 AA340253 F08031 AA192540 H67636 AA321827 AW950283 AA084159 BE538808 AW401377 AA256774 C03366 W46595 W47608 AA305009 H69431 H69456 AL120082 H11706 AA303717 AA361357 H22042 H78020 AW999584 AA134368 AA322811 AA322861 H60980 N85248 N31547 H79624 T11718 W85828 AW894663 AW894624 BE167441 BE170015 AA304626 AW602163 AW998929 AA156681 AA151067 BE002724 AA608688 H82692 BE155392 AW383636 BE155394 AA487004 AW383504 AI342365 R82553 W16498 BE155344 AH143938 R69901 AA322873 AW340648 R25364 AA367935 AI559406 AA033522 AA374252 AW835019 AI922133 AI697089 N99662 AW189078 AI199076 AW151598 W59944 AA662875 W94022 AA299055 AI039008 AI829449 AA583503 AI635674 AW131665 AI473820 AW273118 AW900930 AA908944 AI688035 AW170272 AI082545 AW468176 AI608761 AI082748 AI911682 AI248943 AI831016 AA192465 AI218477 AA938406 AA385288 AI809817 AA905196 AI191245 AI470204 AI188296 AI421367 AI125315 AI087141 AA628032 AA740589 AI554181 AA150830 AI248541 AI077943 AA775958 AA864930 AI261476 AI123121 AI310394 AA862331 AA872478 BE537084 AI205608 AA720684 AI872093 AW150042 AL120538 AA219627 AA986608 C21397 AI593337 H25337 AI069749 AA605146 AI359620 AA150478 AI359738 AW383642 AW995424 AI766457 R56892 AI089839 W61343 N69107 W46459 AA565955 N20527 AI279782 W46596 AA776573 H23204 AI866231 AI083995 N21530 AA126874 D82630 W65437 AI086917 AW382095 AI086877 H69844 AW340217 W85827 L08439 AA262704 AA505380 W47413 W94135 AA223241 AW089153 AA084101 BE538000 AA096126 T28031 AA491574 R84813 AA774538 AW383522 AA155615 AW383529 AA491520 AW028427 AA171456 AI469689 AW664539 AI811102 AI811116 BE464590 BE350791 H78021 T15405 H21979 AA219489 H13301 AA505883 AI864305 AI423963 AW084401 F04963 R69858 H67097 AI917740 AI655561 H69864 AA033631 AW383484 AI866261 H25293 AA513281 AW271187 H11617 N79982 AI174338 AI904207 AI904208 BE614558 W94127 W65436 AI272449 AA700018 AI579932 AI085941 AW152629 AA334551 BE008229 AA307537 AW961156 AW995894 AW995826 NM_006751 M61199 AA045603 AL036372 AV645608 AI688095 AW351901 AA101337 AA101345 N73342 BE018030 BE569044 AW841975 AA373388 BE090412 H95440 N53845 R67867 AA093441 AA363427 H93708 AW023134 AW994986 AW994989 BE090429 R23614 AI567932 H03726 H01101 H01867 AA548743 AI671808 AW872949 AW872941 AA742447 AI189788 AA045604 AI637465 AI741796 AW242217 AW131463 AI765302 AI683923 AA889762 AI804889 AI966437 C06049 BE502340 AI695651 AI491970 AA496804 AA281008 AA665699 AI473814 BE301445 AA707837 AA551925 AI017348 AI208185 AA775203 AA156296 AA557463 H95441 AA768547 AW769358 AA991197 AA181954 AI091369 AI147289 AW771837 AI638582 AA844411 AI374750 T29320 AW951272 AW085923 H02834 AA843259 AA814696 AW183290 AA158453 N68125 N69039 AA100423 AA101346 AI918720 H01102 R67868 H01868 N66438 R46580 AI858433 AA599560 AA187577 AA157481 AA361520 AL047827 AA158452 R21688 AW964874 AA325161 R40871 AW752395 AW375924 R13355 AA281174 AA428908 AA081258 AA160311 W17034 H83596 Z99393 AI831206 AW771108 AW769214 N89775 AW161495 AW161522 AW160880 Z99394 AI814820 BE244667 BE241813 BE242271 AA381943 NM_016040 AF151858 AW967497 AW966873 AI824386 AW470133 AW015765 BE018650 AW503659 AI129638 AI632346 AA013099 AW770511 BE219482 AI824135 AI867379 AA019348 AA285143 AW087624 AI950100 AA251084 AI633962 AA287714 AA400773 AI292112 AW690905 AA743312 AW117423 AA694551 AA885657 AA112675 BE327333 AA082161 H03613 AA094735 AW500235 N28878 AA287713 AW300233 AA826249 AW6921 BE348728 AW505056 AW966879 AI521202 AA393405 AI264668 AA910851 AA251721 AI470834 H03503 AA089688 R58562 BE004728 AA668793 H27167 R54717 W02363 N80298 AA304486 AW954799 AW805136 AW970817 AW373398 AW875459 AA136805 AA683501 N73299 AW341082 AI632954 AA493369 AI478433 AI037911 AW272169 AW043832 AA010683 AW629090 AW183622 N64510 AW079953 AI554533 AA563670 AA010682 AW237610 AW419057 AI470926 AI627833 AA195080 AA195179 AW741443 AW590266 AI168477 AW771214 AI767341 AW340086 AW748455 AI280079 AI244821 AI381283 AW300130 AW183374 AW195397 AA136706 AI24598 AW573004 Z98448 AA905255 AI497883 AW450979 AA136653 AA136656 AW419381 AA984358 AA492073 BE168945 AA809054 AW238038 BE011212 BE011359 BE011367 BE011368 BE011362 BE011215 BE011365 BE011363 AA376654 W76367 AA318232 AI694545 AI742403 AI887383 AW204731 AW874431 BE220997 AA114979 AA303838 AI002267 AW952031 W74801 AA011287 AA115112 AI306385 R37677 AW571707 R59986 W94102 AW197042 H10206 AW139819 AI686172 AI674165 R51633 AI367086 T23948 H10833 H23002 H11743 R37085 Z39208 H22794 H11820 R13817 Z43122 H10257 R88398 R18795 AA010848 R67191 H10875 R67170 AA179656 AA182626 AA182603 R69751 R70467 H69771 H80879 H80878 Z41572 R39330 AA402505 AA398900 X94703 NM_004249 R52316 T87420 N46403 Z36855 BE076834 AW602528 BE073859 Z38412 AA404418 AI217248 AA321890 R18000 AI453076 AI376075 AI014836 AA628633 AA961066 AI150282 AI028574 AI217182 AA732910 AA431478 AL041229 AA353093 AW957317 AW872498 AI560785 AI289110 AW135512 X97261 T68873 AF008937 AI743860 N49543 AW027759 BE349467 AI656284 BE463975 R35022 AA370031 AW955302 AL042109 N53092 AI611424 AL079382 AI969290 AI928016 BE394912 BE504220 BE467505 AI611611 AI611407 AI611452 W56437 AI284566 AI583349 AW183058 AI308085 AI074952 AA437315 AA628161 AW301728 AI150224 AA400137 AA437279 AI223355 AA639462 AI261373 AI432414 AI964994 AI539335 AA401550 AA356757 AI609976 AA442357 AA359393 AA437046 AA370301 AA429328 AW272055 AI580502 AI832944 AI038530 AA425107 AI014986 AI148349 AW237721 AW779756 AW137877 AI125293 AA400404 R28554 AA065069 AA085108 AA069818 AA069971 AA069923 AA069908
107832	genbank_AA021473	AA021473
123523	genbank_AA608588	AA608588
123533	genbank_AA608751	AA608751
132225	genbank_AA128980	AA128980
125017	genbank_T68875	T68875
125063	genbank_T85352	T85352
125064	genbank_T85373	T85373

	125091	genbank_T91518	T91518	
	100964	entrez_J00212	J00212	
	102269	entrez_U30245	U30245	
	125150	NOT_FOUND_entrez_W38240	W38240	
5	123964	genbank_C13961	C13961	
	118111	genbank_N55493	N55493	
	118129	genbank_N57493	N57493	
	102491	entrez_U51010	U51010	
10	118329	genbank_N63520	N63520	
	118475	genbank_N66845	N66845	
	118581	genbank_N68905	N68905	
	111514	genbank_R07998	R07998	
	104534	R22303_at	R22303	
15	120340	genbank_AA206828	AA206828	
	120376	genbank_AA227469	AA227469	
	104787	genbank_AA027317	AA027317	
	120409	genbank_AA235050	AA235050	
	120745	genbank_AA302809	AA302809	
	120809	genbank_AA346495	AA346495	
20	120839	genbank_AA348913	AA348913	
	113702	genbank_T97307	T97307	
	106596	304084_1	AI583948 AA578212 AW303715 AA653450 AA456981 AJ400385 W88533 AI224133 AW272145 AA088686 R94698	
	113947	genbank_W84758	W84758	
25	122562	genbank_AA452156	AA452156	
	122635	genbank_AA454085	AA454085	
	108277	genbank_AA064859	AA064859	
	108403	genbank_AA075374	AA075374	
	122860	genbank_AA464414	AA464414	
30	108427	genbank_AA076382	AA076382	
	108439	genbank_AA078986	AA078986	
	131353	231290_1	AW411259 H23555 AW015049 AI684275 AW015886 AW068953 AW014085 AI027260 R52686 AA818278 AI129462	
			AA89360 N34869 AI948416 AA534205 AA702483 AA705292	
	108533	genbank_AA084415	AA084415	
35	124254	genbank_H69899	H69899	
	101447	entrez_M21305	M21305	
	101458	entrez_M22092	M22092	
	101667	13349_1	NM_005381 M60858 AW373732 AW373724 AW373689 AW373629 AW373609 AW373776 AA187806 AW386946	
40			AW374207 T05235 AA216203 AW385556 AA306940 AA306526 AA315461 AL036757 AW373711 AW403124 AW403640	
			AW377084 T27360 H62836 F06957 AW377051 AA554779 AA378568 AA096007 AW352407 AW302637 F07929 H17433	
			AW382712 H05665 F07292 N39875 AA089729 H62556 N42842 R12952 AW373735 AW364155 AA056183 W39185	
			AW382708 N32488 AF114096 AW375993 AI133569 W52561 AA603040 AA133710 AI928796 AW176370 AA827519	
			AW338437 AA521142 T29341 AI800481 AW317002 AA703914 AA860830 AI859203 AI445772 AA714334 AI817086	
			AI832027 AW510442 AI835802 AW088306 AW088672 AW408555 AW467542 AA552657 AA152367 W32081 AA582124	
45			AA074040 AA931657 AJ051154 AW410203 AI921644 H17434 AI832330 AW404836 AI925038 AA088423 AA954166	
			AA580453 AW021292 AI267215 AW080082 AW383778 AI933053 AI919097 W31557 N80245 AA931591 AA563995	
			F36352 AA056184 AA476294 AA641327 AA533550 AI749630 W58323 AA569119 AA508573 AI809050 AI378998	
			AA411362 AW407505 AA938104 AA074041 AA632876 AW193748 AA507873 AI270128 AI472365 AA411363 AI523216	
			AI719965 AI816302 AA182681 AI707990 AA133588 AI758537 W60253 AI460308 AA135423 AI083904 F04188 N89693	
50			AW408776 AI678595 AI270568 AA722059 W58234 F33650 AA090547 AA285108 AA425981 N85079 D20218 AI273980	
			AA159028 F03226 AW247914 N26918 AW272741 N90109 H05668 N23327 AW247953 R44748 AA62015 F03558	
			AI752394 AW409913 AW248398 AI816463 AI752393 AA325370 AA263089 AI570130 AI971951 AI160658 AI357360	
			AW168688 AL121075 AW050538 N21872 W67748 AA514242 AI127386 H14607 AI185752 W79364 AA088520 AA152476	
			AW351940 AW373683 AI940524 AW374953 T56500 N24329 AI940720 AW374933 AW374947 AW391813 AL138337	
55			AW376241 AW062943 F26686 AW410202 AW062958 F34529 AW381807 AW393315 W17147 AW176359 AA664576	
			AW380424 AA306040 AI745674 AW300951 AI188579 AI438973 AI305271 AA433818 AA612807 AI831809 AI940409	
			AA158663 AI572988	
60	124576	genbank_N68201	N68201	
	108931	genbank_AA147186	AA147186	
	108941	genbank_AA148650	AA148650	
	124720	144582_1	R05283 R11056	
	124793	genbank_R44519	R44519	
	124799	genbank_R45088	R45088	
	103138	entrez_X65965	X65965	
65	117683	genbank_N40180	N40180	
	124991	genbank_T50116	T50116	
	103432	entrez_X97748	X97748	
	119174	genbank_R71234	R71234	
	119239	95573_2	T11483 T11472	
70	133678	11235_1	AW247252 AA346143 NM_000270 AA381085 N91995 X00737 AA381079 AA296473 AA296110 AA315735 AA311617	
			AA326750 AA376804 AW403290 T95231 M13953 T47963 H82039 AA279899 AA627997 N76320 N99527 H37842	
			W20095 AA457308 AW469547 AA724143 H83220 AA319496 W86334 W30892 R89169 R99427 N41854 H47286	
			AA348094 AA045089 R63016 AI922219 AI024906 AI096488 AI885005 AA194872 N90489 AA452544 H72411 AA282427	
			AA430735 R68963 R22453 H70385 AW129369 AW467320 AW519082 AA345018 AA582183 AI981789 R65918 N30611	
			AI979189 AI280889 AW273191 R66531 AI25845 AI875927 AI421990 AW190879 H37794 AA699667 H68427 AA954388	
75			AI188757 AI140048 AA430382 AI204151 AW247864 AA559099 AH31420 AA548276 AI149466 AA772669 AA694388	
			AA724168 AA301651 AA281952 AA779925 AA234760 W86290 AA913603 AW511745 AI500697 AA814922 AA835040	

WO 02/079492

PCT/US02/04915

5 T47964 H53998 AA975804 R98710 AI077604 N70252 R98084 AW250171 H69268 AI597614 AA970746 AA972548
 AI377116 R62962 H16737 R89070 AA731329 R66532 N54354 AI818832 H81944 N71567 T95122 W86463 AA437095
 AI431999 AI915724 N83851 AI674743 AA457307 AA211475 N84444 AI799146 H72853 R99335 H60413 AA770367
 AA156105 AI269937 H64029 H89728 R65819 AW470496 AI873318 AI735713 H82987 C02447 AI478666 T27651
 AI699770 AW025156 H69719 AI984717 N69225 AI459856 AA953577 AI424691 H13843 R22404 AI873796 AI336002
 N70898 AI420854 AA541792 AA346142 AI000814 AI828348 AA045090 T51257 N90434 H13890 N73184 AI708083
 AA781606 AA329050 AA339965 R68964 H64795 W04186 H16845

10 119416 genbank_T97186 T97186
 119558 NOT_FOUND_entrez_W38194 W38194
 119569 NOT_FOUND_entrez_W38197 W38197
 119654 genbank_W57759 W57759

15 121350 genbank_AA405237 AA405237
 121558 genbank_AA412497 AA412497
 105985 genbank_AA406610 AA406610
 114648 genbank_AA101056 AA101056
 121895 genbank_AA427396 AA427396
 100327 entrez_D55640 D55640
 123315 714071_1 AA496369 AA496646
 123473 genbank_AA599143 AA599143

TABLE 4:

5	Pkey:	Unique Eos probeset identifier number			
	Accession:	Accession number used for previous patent filings			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
10	Pkey	Accession	ExAccn	UniGene	UnigeneTitle
15	100405	D86425	AW291587	Hs.82733	nitrogen 2
	100420	D86983	D86983	Hs.118893	Melanoma associated gene
	100481	HG1098-HT1098	X70377	Hs.121489	cystatin D
	100484	HG1103-HT1103	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene hom
	100718	HG3342-HT3519	BE295928	Hs.75424	inhibitor of DNA binding 1, dominant neg
20	100991	J03764	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	101097	L06797	BE245301	Hs.89414	chemokine (C-X-C motif), receptor 4 (fus
	101168	L15388	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	101194	L20971	L20971	Hs.188	phosphodiesterase 4B, cAMP-specific (dun
	101261	L35545	D30857	Hs.82353	protein C receptor, endothelial (EPCR)
25	101345	L76380	NM_005795	Hs.152175	calcitonin receptor-like
	101447	M21305	M21305		gbHuman alpha satellite and satellite 3
	101485	M24736	AA296520	Hs.89546	selectin E (endothelial adhesion molecu
	101543	M31166	M31166	Hs.2050	pentaxin-related gene, rapidly induced b
	101550	M31551	Y00630	Hs.75716	serine (or cysteine) proteinase inhibito
30	101560	M32334	AW958272	Hs.347326	intercellular adhesion molecule 2
	101674	M61916	NM_002291	Hs.82124	laminin, beta 1
	101714	M68874	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
	101741	M74719	NM_003199	Hs.326198	transcription factor 4
	101838	M92934	BE243845	Hs.75511	connective tissue growth factor
35	101857	M94856	BE550723	Hs.153179	fatty acid binding protein 5 (psorlase-
	102012	U03057	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
	102024	U03877	AA301867	Hs.76224	EGF-containing fibulin-like extracellular
	102164	U18300	NM_000107	Hs.77602	damage-specific DNA binding protein 2 (4
	102241	U27109	NM_007351	Hs.268107	multimerin
40	102283	U31384	AW161552	Hs.83381	guanine nucleotide binding protein 11
	102303	U33053	U33053	Hs.2499	protein kinase C-like 1
	102564	U59423	U59423	Hs.79067	MAD (mothers against decapentaplegic, Or
	102663	U70322	NM_002270	Hs.168075	karyopherin (importin) beta 2
	102759	U81607	NM_005100	Hs.788	A kinase (PRKA) anchor protein (gravin)
45	102778	U83463	AF000652	Hs.8180	syndecan binding protein (syntenin)
	102804	U89942	NM_002318	Hs.83354	lysyl oxidase-like 2
	102887	X04729	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	102898	X06256	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,
	102915	X07820	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
50	103036	X54925	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	103037	X54936	BE018302	Hs.2894	placental growth factor, vascular endoth
	103095	X60957	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and
	103158	X67235	BE242587	Hs.118651	hematopoietically expressed homeobox
	103166	X67951	AA159248	Hs.180909	peroxiredoxin 1
55	103185	X69910	NM_006825	Hs.74368	transmembrane protein (63kD), endoplasm
	103280	X79981	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	103554	Z18951	AJ878826	Hs.74034	caveolin 1, caveolae protein, 22kD
	103850	AA187101	AA187101	Hs.213194	hypothetical protein MGC10895
	104465	N24990	Z44203	Hs.28418	ESTs
60	104592	R81003	AW630488	Hs.25338	protease, serine, 23
	104764	AA025351	AJ039243	Hs.278585	ESTs
	104786	AA027168	AA027167	Hs.10031	KIAA0955 protein
	104850	AA040465	AL133035	Hs.8728	hypothetical protein DKFZp434G171
	104865	AA045136	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
65	104894	AA054087	AF065214	Hs.18858	phospholipase A2, group IVC (cytosolic,
	104952	AA071089	AW076098	Hs.345588	desmoplakin (DPI, DP1)
	104974	AA085918	Y12059	Hs.278675	bromodomain-containing 4
	105178	AA187490	AA313825	Hs.21941	AD036 protein
	105263	AA227926	AW388633	Hs.6682	solute carrier family 7, (cationic amino
70	105330	AA234743	AW338625	Hs.22120	ESTs
	105376	AA236559	AW994032	Hs.8768	hypothetical protein FLJ10849
	105729	AA282694	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	105826	AA398243	AA478756	Hs.194477	E3 ubiquitin ligase SMURF2
	105977	AA406363	AK001972	Hs.30822	hypothetical protein FLJ11110
75	106008	AA411465	AB033888	Hs.8619	SRY (sex determining region Y)-box 18
	106031	AA412284	X64116	Hs.171844	Homo sapiens cDNA: FLJ22296 fls, clone H
	106124	AA423987	H93366	Hs.7567	Homo sapiens cDNA: FLJ21962 fls, clone H

	106155	AA425309	AA425414	Hs.33287	nuclear factor I/B
	106302	AA435896	AA398859	Hs.18397	hypothetical protein FLJ23221
	106423	AA448238	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15
5	106793	AA478778	H94997	Hs.18450	ESTs
	107174	AA621714	BE122782	Hs.25338	ESTs
	107216	D51069	D51069	Hs.211579	melanoma cell adhesion molecule
	107295	T34527	AA186629	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
	107385	U97519	NM_005397	Hs.16426	podocalyxin-like
10	108756	AA127221	AA127221	Hs.117037	ESTs
	108846	AA132983	AL117452	Hs.44155	DKFZP586G1517 protein
	108888	AA135606	AA135606	Hs.189384	gb:z110a05.s1 Soares_pregnant_uterus_NbH
	109001	AA156125	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to
	109166	AA179845	AA218691	Hs.73625	RAB6 interacting, kinesin-like (rabkines
15	109456	AA232645	AW956580	Hs.42699	ESTs
	109768	F10399	F06838	Hs.14763	ESTs
	110107	H16772	AW151660	Hs.31444	ESTs
	110906	N39584	AA035211	Hs.17404	ESTs
	110984	N52006	AW613287	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
20	111006	N53375	BE387014	Hs.166146	Homer, neuronal immediate early gene, 3
	111018	N54087	AI287912	Hs.3628	mitogen-activated protein kinase kinase
	111133	N64436	AW508939	Hs.97189	complement component C1q receptor
	111760	R26892	BE551929	Hs.268754	Homo sapiens cDNA FLJ11949 fls, clone HE
	113073	T33637	N39342	Hs.103042	microtubule-associated protein 1B
25	113195	T57112	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	113923	W80763	AW953484	Hs.3849	hypothetical protein FLJ22041 similar to
	114521	AA046808	AW139036	Hs.108957	40S ribosomal protein S27 isoform
	115061	AA253217	AI751438	Hs.41271	Homo sapiens mRNA full length insert cDN
	115096	AA255991	AI683069	Hs.175319	ESTs
30	115145	AA268138	AA740907	Hs.88297	ESTs
	115819	AA426573	AA486620	Hs.41135	endomucin-2
	115947	AA443793	R47479	Hs.94761	KIAA1691 protein
	116314	AA490588	AI799104	Hs.178705	Homo sapiens cDNA FLJ11333 fls, clone PL
	116339	AA496257	AK000290	Hs.44033	dipeptidyl peptidase 8
35	116430	AA609717	AK001531	Hs.66048	hypothetical protein FLJ10669
	116589	D59570	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	116733	F13787	AL157424	Hs.61289	synaptotagmin 2
	117023	H88157	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
	117186	H98988	H98988	Hs.42612	ESTs, Weakly similar to ALU1_HUMAN ALU S
40	117563	N34287	AF055634	Hs.44553	unc5 (C.elegans homolog) c
	117997	N52090	N52090	Hs.47420	EST
	118475	N66845	N66845		gb:za46c11.s1 Soares fetal liver spleen
	118581	N68905	N68905		gb:za65b09.s1 Soares_fetal_lung_NbHL19W
	119073	R32894	BE245360	Hs.279477	ESTs
45	119155	R61715	R61715	Hs.310598	ESTs, Moderately similar to ALU1_HUMAN A
	119174	R71234	R71234		gb:yl54c08.s1 Soares placenta Nb2HP Homo
	119221	R98105	C14322	Hs.250700	tryptase beta 1
	119416	T97186	T97186		gb:ye50h09.s1 Soares fetal liver spleen
50	119668	W80814	AA496205	Hs.193700	Homo sapiens mRNA; cDNA DKFZp586I0324 (f
	121335	AA404418	AA404418		gb:zw37e02.s1 Soares_total_fetus_Nb2HF8_
	121381	AA405747	AW088642	Hs.97984	hypothetical protein FLJ22252 similar to
	123160	AA486687	AA486687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti
	123473	AA599143	AA599143		gb:ae52d04.s1 Stratagene lung carcinoma
	123523	AA608588	AA608588		gb:ae54e06.s1 Stratagene lung carcinoma
55	123533	AA608751	AA608751		gb:ae56h07.s1 Stratagene lung carcinoma
	123964	C13961	C13961		gb:C13961 Clontech human aorta polyA+ mR
	124006	D60302	AI147155	Hs.270016	ESTs
	124315	H94892	NM_005402	Hs.288757	v-rat simian leukemia viral oncogene hom
	124659	N93521	AI690737	Hs.289068	Homo sapiens cDNA FLJ11918 fls, clone HE
60	124669	N95477	AI571594	Hs.102943	hypothetical protein MGC12916
	124847	R60044	W07701	Hs.304177	Homo sapiens clone FLB8503 PROZ286 mRNA,
	124875	R70506	AI887664	Hs.285814	sprouty (Drosophila) homolog 4
	125091	T91518	T91518		gb:ye20f05.s1 Stratagene lung (937210) H
	125103	T95333	AA570056	Hs.122730	ESTs, Moderately similar to KIAA1215 pro
65	125355	R45630	R60547	Hs.170098	KIAA0372 gene product
	125565	R20839	R20840		gb:yg05c08.r1 Soares Infant brain 1NIB H
	125590	R23858	R23858	Hs.143375	Homo sapiens, clone IMAGE:3840937, mRNA,
	125765	R23858	R23858	Hs.143375	Homo sapiens, clone IMAGE:3840937, mRNA,
	126511	AI024874	T92143	Hs.57958	EGF-TM7-latrophilin-related protein
70	100286	W26247	BE247550	Hs.86859	growth factor receptor-bound protein 7
	126563	W26247	AA516391	Hs.181368	U5 snRNP-specific protein (220 kD), orth
	126649	AA856990	AA001860	Hs.279531	ESTs
	126672	AA136653	AA001860	Hs.279531	ESTs
	126872	AA136653	AW450979		gb:UH-H-BI3-ala-a-12-0-UL.s1 NCL_CGAP_Su
75	456000	AA136653	BE180878	Hs.11614	HSPC065 protein
	141221	AA136653	AW450979		gb:UH-H-BI3-ala-a-12-0-UL.s1 NCL_CGAP_Su
	127402	AA358869	AA358869	Hs.227949	SEC13 (S. cerevisiae)-like 1

	127651	AI123976	AA382523	Hs.105689	MSTP031 protein
	424806	AI123976	AA382523	Hs.105689	MSTP031 protein
	128082	AA379500	AA379521	Hs.105547	neural proliferation, differentiation an
5	128992	R49693	H04150	Hs.107708	ESTs
	129046	AA195678	AB029290	Hs.108258	actin binding protein; macrophin (microf
	129188	M30257	NM_001078	Hs.109225	vascular cell adhesion molecule 1
	129314	AA028131	BE622768	Hs.290356	mesoderm development candidate 1
	129371	M10321	X06828	Hs.110802	von Willebrand factor
10	129468	J03040	AW410538	Hs.111779	secreted protein, acidic, cysteine-rich
	129765	M86933	M86933	Hs.1238	amelogenin (Y chromosome)
	129805	AA012933	AA012848	Hs.12570	tubulin-specific chaperone d
	129884	AA286710	AF065581	Hs.13131	lysosomal
	130495	AA243278	AW250380	Hs.109059	mitochondrial ribosomal protein L12
15	130639	D59711	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	130657	T94452	AW337575	Hs.201591	ESTs
	130828	AA053400	AW631469	Hs.203213	ESTs
	130972	AA370302	D81866	Hs.21739	Homo sapiens mRNA; cDNA DKFZp586i1518 (f
	131080	J05008	NM_001955	Hs.2271	endothelin 1
20	131137	U85193	W27392	Hs.33287	nuclear factor I/B
	131182	AA256153	AI824144	Hs.23912	ESTs
	131486	X83107	F06972	Hs.27372	BMX non-receptor tyrosine kinase
	131573	AA046593	AA040311	Hs.28959	ESTs
	131647	AA410480	AA359615	Hs.30089	ESTs
25	131756	D45304	AA443966	Hs.31595	ESTs
	131859	M90657	AW960564		transmembrane 4 superfamily member 1
	131881	AA010163	AW361018	Hs.3383	upstream regulatory element binding prot
	132050	AA136353	AI267615	Hs.38022	ESTs
	132083	Y07867	BE386490	Hs.279663	Pirin
30	132164	U84573	AI752235	Hs.41270	procollagen-tyrosine, 2-oxoglutarate 5-dio
	132356	X60486	NM_003542	Hs.46423	H4 histone family, member G
	132413	AA132969	AW361383	Hs.260116	metalloprotease 1 (pitrilysin family)
	132456	AA114250	AB011084	Hs.48924	KIAA0512 gene product; ALEX2
	132490	F13782	NM_001290	Hs.4980	LIM domain binding 2
35	132676	AA283035	N92589	Hs.261038	ESTs, Weakly similar to I38022 hypotheti
	132687	AB002301	AB002301	Hs.54985	KIAA0303 protein
	132718	AA056731	NM_004600	Hs.554	Sjogren syndrome antigen A2 (60kD, ribon
	132736	U68019	AW081883	Hs.211578	Homo sapiens cDNA: FLJ23037 fis, clone L
	132760	H99198	AA125985	Hs.56145	thymosin, beta, identified in neuroblast
40	132933	AA58702	BE263252	Hs.6101	hypothetical protein MGC3178
	132968	N77151	AF234532	Hs.61638	myosin X
	132994	AA505133	AA112748	Hs.279905	clone HQ0310 PRO0310p1
	133081	AB000584	AI186431	Hs.296638	prostate differentiation factor
	133147	D12763	AA026533	Hs.66	interleukin 1 receptor-like 1
45	133161	AA253193	AW021103	Hs.6631	hypothetical protein FLJ20373
	133200	AA432248	AB037715	Hs.183639	hypothetical protein FLJ10210
	133260	AA083572	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone R
	133363	AA479713	AI866286	Hs.71962	ESTs, Weakly similar to B36298 prolina-r
	133491	L40395	BE619053	Hs.170001	eukaryotic translation initiation factor
50	133517	X52947	NM_000165	Hs.74471	gap junction protein, alpha 1, 43kD (con
	133550	W80846	AI129903	Hs.74669	vesicle-associated membrane protein 5 (m
	133607	M34539	BE273749		FK506-binding protein 1A (12kD)
	133614	D67029	NM_003003	Hs.75232	SEC14 (S. cerevisiae)-like 1
	133627	U09587	NM_002047	Hs.75280	glycyl-tRNA synthetase
55	133691	M85289	M85289	Hs.211573	heparan sulfate proteoglycan 2 (perlecan
	133696	D10522	AI878921	Hs.75607	myristoylated alanine-rich protein kinas
	133913	W84712	AU076964	Hs.7753	calumenin
	133975	D29992	C18356	Hs.295944	tissue factor pathway inhibitor 2
	133985	L34657	L34657	Hs.78146	platelet/endothelial cell adhesion molec
60	134039	S78569	NM_002290	Hs.78672	laminin, alpha 4
	134088	D43636	AI379954	Hs.79025	KIAA0096 protein
	134161	U97188	AA634543	Hs.79440	IGF-II mRNA-binding protein 3
	134299	AA487558	AW580939	Hs.97199	complement component C1q receptor
	134416	M28882	X68264	Hs.211579	melanoma cell adhesion molecule
65	134670	X70683	AI272141	Hs.83484	SRY (sex determining region Y)-box 4
	134656	X14787	AI750878	Hs.87409	thrombospondin 1
	134989	AA236324	AW968058	Hs.92381	nudix (nucleoside diphosphate linked mol
	135051	C15324	AI272141	Hs.83484	SRY (sex determining region Y)-box 4
	135073	AA452000	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
70	135349	D83174	AA114212	Hs.9930	serine (or cysteine) proteinase inhibitor
	100114	D00596	X02308	Hs.82962	thymidylate synthetase
	100130	D11428	NM_000304	Hs.103724	peripheral myelin protein 22
	100143	D13640	AU076465	Hs.278441	KIAA0015 gene product
	100168	D14874	H73444	Hs.394	adrenomedullin
75	100208	D26129	NM_002933	Hs.78224	ribonuclease, RNase A family, 1 (pancrea
	100224	D28476	AL121516	Hs.138617	thyroid hormone receptor interactor 12
	100405	D86425	AW291587	Hs.82733	nidogen 2

	100420	D86983	D86983	Hs.118893	Melanoma associated gene
	100455	D87953	AW888941	Hs.75789	N-myc downstream regulated
	100529	HG1862-HT1897	BE313693	Hs.334330	calmodulin 2 (phosphorylase kinase, delt
	100618	HG2614-HT2710	A1752163	Hs.114599	collagen, type VIII, alpha 1
5	100619	HG2639-HT2735	N24433	Hs.241567	RNA binding motif, single stranded Inter
	100658	HG2855-HT2995	U56725	Hs.180414	heat shock 70kD protein 2
	100676	HG3044-HT3742	X02761	Hs.287820	fibronectin 1
	100718	HG3342-HT3519	BE295928	Hs.75424	inhibitor of DNA binding 1, dominant neg
10	100752	HG3543-HT3739	T81309		insulin-like growth factor 2 (somatomedi
	100828	HG4069-HT4339	AL048753	Hs.303649	small inducible cytokine A2 (monocyte ch
	100850	HG417-HT417	AA836472	Hs.297939	cathepsin B
	100991	J03764	J03836	Hs.82085	serine (or cysteine) proteinase inhibito
	101097	L06797	BE245301	Hs.89414	chemokine (C-X-C motif), receptor 4 (fus
15	101110	L08248	A1439011	Hs.86386	myeloid cell leukemia sequence 1 (BCL2-r
	101142	L12711	L12711	Hs.89643	transketolase (Wemicke-Korsakoff syndro
	101156	L13977	AA340987	Hs.75693	prolylcarboxypeptidase (angiotensinase C
	101168	L15388	NM_005308	Hs.211569	G protein-coupled receptor kinase 5
	101184	L19871	NM_001674	Hs.460	activating transcription factor 3
20	101192	L20859	BE247295	Hs.78452	solute carrier family 20 (phosphate tran
	101317	L42176	L42176	Hs.8302	four and a half LIM domains 2
	101336	L49169	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h
	101345	L76380	NM_005795	Hs.152175	calcitonin receptor-like
	101400	M15990	M15990	Hs.194148	v-src-1 Yamaguchi sarcoma viral oncogene
25	101475	M23254	BE410405	Hs.76288	calpain 2, (mII) large subunit
	101485	M24736	AA296520	Hs.89546	selectin E (endothelial adhesion molecul
	101496	M26576	X12784	Hs.119129	collagen, type IV, alpha 1
	101505	M27396	AA307680	Hs.75692	asparagine synthetase
	101543	M31166	M31166	Hs.2050	pentaxin-related gene, rapidly induced b
30	101557	M31994	BE293116	Hs.76392	aldehyde dehydrogenase 1 family, member
	101560	M32334	AW958272	Hs.347326	intercellular adhesion molecule 2
	101587	M35878	A1752416	Hs.77326	insulin-like growth factor binding prote
	101592	M36429	AF064853	Hs.91299	guanine nucleotide binding protein (G pr
	101633	M57730	NM_004428	Hs.1624	ephrin-A1
35	101634	M57731	AV650282	Hs.75765	GRO2 oncogene
	101667	M60858	NM_005381		nucleolin
	101682	M62994	AF043045	Hs.81008	filamin B, beta (actin-binding protein-2
	101714	M68874	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
	101720	M69043	M69043	Hs.81328	nuclear factor of kappa light polypeptid
40	101741	M74719	NM_003199	Hs.326198	transcription factor 4
	101744	M75126	A1879352	Hs.118625	hexokinase 1
	101793	M84349	W01076	Hs.276573	CD59 antigen p18-20 (antigen identified
	101837	M92843	M92843	Hs.343586	zinc finger protein homologous to Zip-36
	101838	M92934	BE243845	Hs.75511	connective tissue growth factor
45	101840	M93056	AA236291	Hs.183583	serine (or cysteine) proteinase inhibito
	101857	M94856	BE550723	Hs.153179	fatty acid binding protein 5 (psoriasis-
	101864	M95787	BE392588	Hs.75777	transgelin
	101931	S76965	NM_006823	Hs.75209	protein kinase (cAMP-dependent, catalyti
50	101966	S81914	X96438	Hs.76095	immediate early response 3
	102012	U03057	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
	102013	U03100	BE616287	Hs.178452	catenin (cadherin-associated protein), a
	102024	U03877	AA301867	Hs.76224	EGF-containing fibulin-like extracellular
	102059	U08021	A1752666	Hs.76669	nicotinamide N-methyltransferase
55	102121	U14391	NM_004998	Hs.82251	myosin IE
	102283	U31384	AW161552	Hs.83381	guanine nucleotide binding protein 11
	102300	U32944	A1929721	Hs.5120	dynein, cytoplasmic, light polypeptide
	102378	U40369	AU076887	Hs.28491	spermidine/spermine N1-acetyltransferase
	102395	U41767	AU077005	Hs.92208	a disintegrin and metalloproteinase doma
60	102460	U48959	U48959	Hs.211582	myosin, light polypeptide kinase
	102491	U51010	U51010		gbHuman nicotinamide N-methyltransferas
	102499	U51478	BE243877	Hs.76941	ATPase, Na ⁺ /K ⁺ transporting, beta 3 poly
	102523	U53445	U53445	Hs.15432	downregulated in ovarian cancer 1
	102560	U59289	R97457	Hs.63984	cadherin 13, H-cadherin (heart)
	102564	U59423	U59423	Hs.79087	MAD (mothers against decapentaplegic, Dr
65	102589	U62015	AU076728	Hs.8867	cysteine-rich, angiogenic inducer, 61
	102600	U63825	A1984144	Hs.66713	hepatitis delta antigen-interacting prot
	102645	U67963	AL119566	Hs.6721	lysosomal
	102687	U73379	NM_007019	Hs.93002	ubiquitin carrier protein E2-C
	102693	U73824	AA532780	Hs.183684	eukaryotic translation initiation factor
70	102709	U77604	AA122237	Hs.81874	microsomal glutathione S-transferase 2
	102759	U81607	NM_005100	Hs.788	A kinase (PRKA) anchor protein (gravin)
	102804	U89942	NM_002318	Hs.83354	lysyl oxidase-like 2
	102882	X04412	A1767736	Hs.290070	gelsoin (amyloidosis, Finnish type)
	102907	X06985	BE409861	Hs.202833	heme oxygenase (decycling) 1
75	102915	X07820	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	102927	X12876	BE512730	Hs.65114	keratin 18
	102960	X15729	A1904738	Hs.76053	DEADH (Asp-Glu-Ala-Asp/His) box polypep

	103011	X52541	AJ243425	Hs.326035	early growth response 1
	103020	X53416	X53416	Hs.195464	filamin A, alpha (actin-binding protein-
	103029	X54489	AW800726	Hs.789	GRO1 oncogene (melanoma growth stimulat
5	103036	X54925	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	103058	X57206	Y18024	Hs.78877	inositol 1,4,5-bisphosphate 3-kinase B
	103080	X59798	AU077231	Hs.82932	cyclin D1 (PRAD1: parathyroid adenomas
	103095	X60957	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and
	103138	X65965	X65965		gb3H.sapiens SOD-2 gene for manganese su
10	103176	X69111	AL021154	Hs.76884	inhibitor of DNA binding 3, dominant neg
	103195	X70940	AA351647	Hs.2642	eukaryotic translation elongation factor
	103347	X87638	AU077309	Hs.171271	catenin (cadherin-associated protein), b
	103371	X91247	X91247	Hs.13046	thioredoxin reductase 1
	103432	X97748	X97748		gb3H.sapiens PTX3 gene promotor region.
15	103471	Y00815	Y00815	Hs.75216	protein tyrosine phosphatase, receptor l
	103967	AA303711	AL120051	Hs.144700	ephrin-B1
	104447	L44538	AW204145	Hs.156044	ESTs
	104764	AA026351	AI039243	Hs.278585	ESTs
	104783	AA027050	AA533513	Hs.93659	protein disulfide isomerase related prot
20	104798	AA029462	AW952619	Hs.17235	Homo sapiens clone TCCCA00176 mRNA sequ
	104865	AA045136	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
	104877	AA047437	AI138635	Hs.22968	Homo sapiens clone IMAGE:451939, mRNA se
	104894	AA054087	AF085214	Hs.18858	phospholipase A2, group IVC (cytosolic,
	104952	AA071089	AW076098	Hs.345588	desmoplakin (DPI, DPL)
25	105113	AA156450	AB037816	Hs.8982	Homo sapiens, clone IMAGE:3506202, mRNA,
	105178	AA187490	AA313825	Hs.21941	AD036 protein
	105196	AA195031	W84893	Hs.9305	angiotensin receptor-like 1
	105215	AA205724	AA205759	Hs.10119	hypothetical protein FLJ14957
	105263	AA227926	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	105271	AA227986	AA807881	Hs.25329	ESTs
30	105330	AA234743	AW338625	Hs.22120	ESTs
	105461	AA253216	BE539071	Hs.69388	hypothetical protein FLJ20505
	105492	AA256210	AI805717	Hs.289112	CGI-43 protein
	105493	AA256268	AL047586	Hs.10283	RNA binding motif protein 8B
35	105594	AA279397	AB024334	Hs.25001	tyrosine 3-monooxygenase/tryptophan 5-mo
	105727	AA292379	AL135159	Hs.20340	KIAA1002 protein
	105732	AA292717	AW504170	Hs.274344	hypothetical protein MGC12942
	105767	AA346551	AW370946	Hs.23457	ESTs
	105882	AA400292	W46802	Hs.81988	disabled (Drosophila) homolog 2 (mitogen
40	105936	AA404338	AI678765	Hs.21812	ESTs
	106031	AA412284	X64116	Hs.171844	Homo sapiens cDNA: FLJ22296 fts, clone H
	106124	AA423987	H93366	Hs.7567	Homo sapiens cDNA: FLJ21962 fts, clone H
	106222	AA428594	AA356392	Hs.21321	Homo sapiens clone FLB9213 PRO2474 mRNA,
	106241	AA430108	BE019681	Hs.6019	Homo sapiens cDNA: FLJ21288 fts, clone C
45	106263	AA431462	W21493	Hs.28329	hypothetical protein FLJ14005
	106264	AA431470	AL046859	Hs.3407	protein kinase (cAMP-dependent, catalyti
	106366	AA443756	AA186715	Hs.336429	RIKEN cDNA 9130422N19 gene
	106454	AA449479	NM_014038	Hs.5216	HSPC028 protein
	106634	AA459916	W25491	Hs.288909	hypothetical protein FLJ22471
50	106724	AA465226	N48670	Hs.28631	Homo sapiens cDNA: FLJ22141 fts, clone H
	106793	AA478778	H94997	Hs.16450	ESTs
	106799	AA479037	BE313412	Hs.7961	Homo sapiens clone 25012 mRNA sequence
	106842	AA482597	AF124251	Hs.26054	novel SH2-containing protein 3
	106868	AA487561	BE185536	Hs.301183	molecule possessing ankyrin repeats indu
55	106890	AA489245	AA489245	Hs.88500	mitogen-activated protein kinase 8 inter
	106961	AA504110	AW243614	Hs.18063	Homo sapiens cDNA FLJ10768 fts, clone NT
	106974	AA520969	AI817130	Hs.9195	Homo sapiens cDNA FLJ13698 fts, clone PL
	107030	AA599434	AL117424	Hs.25035	chloride intracellular channel 4
	107061	AA608649	BE147611	Hs.6354	stromal cell derived factor receptor 1
60	107086	AA609519	NM_012331	Hs.26458	methionine sulfoxide reductase A
	107216	D51069	D51069	Hs.211579	melanoma cell adhesion molecule
	107385	U97519	NM_005397	Hs.16426	podocalyxin-like
	107444	W28391	W28391	Hs.343258	proliferation-associated 2G4, 38kD
	107985	AA035638	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
65	108507	AA083514	AI554545	Hs.68301	ESTs
	108695	AA121315	AB029000	Hs.70823	KIAA1077 protein
	108931	AA147186	AA147186		gbz038d01.s1 Stratagene endothelial cel
	109001	AA156125	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to
	109195	AA188932	AF047033	Hs.132904	solute carrier family 4, sodium bicarbon
70	109390	AA219653	AW007485	Hs.87125	EH-domain containing 3
	109456	AA232645	AW956580	Hs.42699	ESTs
	109737	F10078	AA055415	Hs.13233	ESTs, Moderately similar to A47582 B-cel
	110411	H48032	AW001579	Hs.9645	Homo sapiens mRNA for KIAA1741 protein,
	110660	H82117	AA782114	Hs.28043	ESTs
	110905	N39584	AA035211	Hs.17404	ESTs
75	111018	N54067	AI287912	Hs.3628	mitogen-activated protein kinase kinase
	111091	N59858	AA300087	Hs.33032	hypothetical protein DKFZp434N185

	111356	N90933	BE301871	Hs.4887	mannosyl (alpha-1,3)-glycoprotein beta-
	111378	N93764	AW160993	Hs.326292	hypothetical gene DKFZp434A1114
	111741	R26124	AB020653	Hs.24024	KIAA0846 protein
	111769	R27957	AW629414	Hs.24230	ESTs
5	112318	R55470	AW083384	Hs.11067	ESTs, Highly similar to T46395 hypotheti
	112951	T16550	AA307634	Hs.6650	vacuolar protein sorting 45B (yeast homo
	113057	T26674	AW194301	Hs.339283	Human DNA sequence from clone RP1-187J11
	113195	T57112	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	113490	T88700	BE178110	Hs.173374	Homo sapiens cDNA FLJ10500 fts, clone NT
10	113542	T90527	H43374	Hs.7890	Homo sapiens mRNA for KIAA1671 protein,
	113803	W42789	AW880709	Hs.283683	chromosome 8 open reading frame 4
	113847	W60002	NM_005032	Hs.4114	plastin 3 (T isoform)
	113910	W78175	AA113262	Hs.17901	Homo sapiens, clone IMAGE:3937015, mRNA,
	113947	W84768	W84768		gbzh53d03.s1 Soares_fetal_liver_spleen_
15	114047	W94427	AL035858	Hs.3807	FXD domain-containing ion transport reg
	115081	AA253217	AI751438	Hs.41271	Homo sapiens mRNA full length insert cDN
	115819	AA426573	AA488620	Hs.41135	endomucin-2
	115870	AA432374	NM_005985	Hs.48029	snail 1 (drosophila homolog), zinc finger
	115964	AA446622	AA887568	Hs.74313	KIAA1265 protein
20	116228	AA478771	AI767947	Hs.50841	ESTs
	116264	AA482594	D51174	Hs.272239	lysosomal
	116314	AA490588	AI799104	Hs.178705	Homo sapiens cDNA FLJ11333 fts, clone PL
	116589	D59570	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	117023	H88157	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
25	117112	H94648	AW969999	Hs.293658	ESTs
	117156	H97538	W73853		ESTs
	117176	H98670	H45100	Hs.49753	uveal autoantigen with coiled coil domai
	117280	N22107	M18217	Hs.172129	Homo sapiens cDNA: FLJ21409 fts, clone C
	119559	W38197	W38197		Empirically selected from AFFX single pr
30	119866	W80814	AA496205	Hs.193700	Homo sapiens mRNA; cDNA DKFZp586N0324 (f
	120655	AA287347	AA305599	Hs.238205	hypothetical protein PRO2013
	121314	AA402799	W07343	Hs.182538	phospholipid scramblase 4
	121335	AA404418	AA404418		gbzw37e02.s1 Soares_fetal_liver_spleen_
35	121822	AA425107	AI743860		metallothionein 1E (functional)
	121835	AA425435	AB033030	Hs.300670	KIAA1204 protein
	122331	AA442872	AL133437	Hs.110771	Homo sapiens cDNA: FLJ21904 fts, clone H
	122577	AA452860	AA829725	Hs.334437	hypothetical protein MGC4248
	123160	AA488687	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti
40	123488	AA599674	BE019072	Hs.334802	Homo sapiens cDNA FLJ14680 fts, clone NT
	124359	F13673	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti
	124339	H99093	H99093	Hs.343411	DEAD(H) (Asp-Glu-Ala-Asp/His) box polypep
	124358	N22495	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f
	124364	N23031	AF265555	Hs.250646	baculoviral IAP repeat-containing 6
	124726	R15740	NM_003654	Hs.104576	carbohydrate (keratan sulfate Gal-6) sul
45	124763	R39810	BE410405	Hs.76288	calpain 2, (mII) large subunit
	125167	W45560	AL137540	Hs.102541	netrin 4
	125304	Z39833	AL359573	Hs.124940	GTP-binding protein
	125307	Z40583	AW580945	Hs.330466	ESTs
50	125329	AA825437	AA825437	Hs.58875	ESTs
	107985	R66613	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
	125598	R66613	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
	125609	AA868063	AA868063	Hs.104576	carbohydrate (keratan sulfate Gal-6) sul
	116024	AA128075	AA088767	Hs.83883	transmembrane, prostate androgen induced
55	418000	AA128075	AA932794	Hs.83147	guanine nucleotide binding protein-like
	126399	AA128075	AA088767	Hs.83883	transmembrane, prostate androgen induced
	127435	N66570	X69086	Hs.286161	Homo sapiens cDNA FLJ13613 fts, clone PL
	127566	AI051390	AI051390	Hs.116731	ESTs
	127619	AA627122	AA627122	Hs.163787	ESTs
	434190	AA627122	AA627122	Hs.163787	ESTs
60	128453	X02761	X02761	Hs.287820	fibronectin 1
	128495	AF010193	NM_005904	Hs.100602	MAD (mothers against decapentaplegic, Dr
	128515	AA149044	BE395085	Hs.10086	type I transmembrane protein Fn14
	128580	U82108	U82108	Hs.101813	solute carrier family 9 (sodium/hydrogen
65	128623	D78676	BE076608	Hs.105509	CTL2 gene
	128642	L35240	Z28913	Hs.102948	enigma (LIM domain protein)
	128669	AA598737	W28493	Hs.180414	heat shock 70kD protein 8
	128903	R69417	AW150717	Hs.345728	STAT induced STAT inhibitor 3
	128914	AA232837	AW867491	Hs.107125	plasmalemma vesicle associated protein
70	129087	N72695	AI348027	Hs.108557	hypothetical protein PP1057
	129188	M30257	NM_001078	Hs.109225	vascular cell adhesion molecule 1
	129226	M96843	BE222494	Hs.180919	inhibitor of DNA binding 2, dominant neg
	129285	X68277	AA530892	Hs.171695	dual specificity phosphatase 1
	129345	AA292440	R22497	Hs.110571	growth arrest and DNA-damage-inducible,
	129468	J03040	AW410538	Hs.111779	secreted protein, acidic, cysteine-rich
75	129488	AA228107	AW966728	Hs.54642	methionine adenosyltransferase II, beta
	101838	AA449789	BE243845	Hs.75511	connective tissue growth factor

	413731	AA449789	BE243845	Hs.75511	connective tissue growth factor
	129557	W01367	AL045404	Hs.46366	KIAA0948 protein
	129619	AA610116	AA209534	Hs.284243	tetraspan NET-6 protein
	129627	AA258308	T40064	Hs.71968	Homo sapiens mRNA; cDNA DKFZp564F053 (fr
5	129762	AA460273	AA453694	Hs.12372	tripartite motif protein TRIM2
	129884	AA286710	AF055581	Hs.13131	lysosomal
	130018	T68873	AA353093		metallothionein 1L
	130147	D63476	D63476	Hs.172813	PAK-interacting exchange factor beta
	130178	M62403	U20982	Hs.1516	insulin-like growth factor-binding prote
10	130282	X55740	BE245380	Hs.153952	5' nucleotidase (CD73)
	130431	L10284	AW505214	Hs.155560	calnexin
	130495	AA243278	AW250380	Hs.109059	mitochondrial ribosomal protein L12
	130553	AA430032	AF082649	Hs.252587	pituitary tumor-transforming 1
	130838	H16402	AW021276	Hs.17121	ESTs
15	130839	D59711	AI557212	Hs.17132	ESTs, Moderately similar to I54374 gene
	130657	T94452	AW337575	Hs.201591	ESTs
	130686	AA431571	BE548267	Hs.337986	Homo sapiens cDNA FLJ10934 fls, clone OV
	130776	R79356	AF167706	Hs.19280	cysteine-rich motor neuron 1
20	130818	AA280375	AW190920	Hs.19928	hypothetical protein SP329
	130840	Z49269	BE048821	Hs.20144	small inducible cytokine subfamily A (Cy
	130899	Z41740	AI077288	Hs.296323	serum/glucocorticoid regulated kinase
	131002	AA121543	AL050295	Hs.22039	KIAA0758 protein
	131080	J05008	NM_001955Hs.2271		endothelin 1
25	131084	AA101878	NM_017413Hs.303084		apelin; peptide ligand for APJ receptor
	131091	T35341	AJ271216	Hs.22880	dipeptidylpeptidase III
	131107	N87590	BE620886	Hs.75354	GCR1 (general control of amino-acid synt
	131182	AA256153	AI824144	Hs.23912	ESTs
	131207	W74533	AF104266	Hs.24212	latrophilin
30	131319	U25997	NM_003155Hs.25590		stannocalcin 1
	131328	V01512	AW939251	Hs.25647	v-fos FBJ murine osteosarcoma viral onco
	131509	X56681	X56681	Hs.2780	jun D proto-oncogene
	131555	AA161292	T47364	Hs.278613	interferon, alpha-inducible protein 27
	131564	AA491465	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fls, clone PL
35	131573	AA046593	AA040311	Hs.28959	ESTs
	131692	D50914	BE559681	Hs.30736	KIAA0124 protein
	131756	D45304	AA443966	Hs.31595	ESTs
	131859	M90657	AW960564		transmembrane 4 superfamily member 1
	131909	W69127	NM_016558Hs.274411		SCAN domain-containing 1
40	131915	AA316186	AI161383	Hs.34549	ESTs, Highly similar to S94541 1 clone 4
	132046	AA384503	AI359214	Hs.179260	chromosome 14 open reading frame 4
	132050	AA136353	AI267615	Hs.38022	ESTs
	132151	AA044755	BE379499	Hs.173705	Homo sapiens cDNA: FLJ22050 fls, clone H
	132164	U84573	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio
45	132187	AA058911	AA235709	Hs.4193	DKFZP586O1624 protein
	132203	AA620962	BE177330	Hs.325093	Homo sapiens cDNA: FLJ21210 fls, clone C
	132314	AA285290	AF112222	Hs.323806	plnin, desmosome associated protein
	132358	X60486	NM_003542Hs.46423		H4 histone family, member G
	132398	R31641	AA876616	Hs.16979	ESTs, Weakly similar to A43932 mucin 2 p
50	132421	AA489190	AW163483	Hs.48320	double ring-finger protein, Dorrin
	132490	F13782	NM_001290Hs.4980		LIM domain binding 2
	132520	AA257993	AA257992	Hs.50651	Janus kinase 1 (a protein tyrosine kinas
	132546	M24283	M24283	Hs.168383	intercellular adhesion molecule 1 (CD54)
	132610	AA443114	AA160511	Hs.5326	amino acid system N transporter 2; porcu
55	132716	T35289	BE379595	Hs.283738	casein kinase 1, alpha 1
	132840	N23817	BE218319	Hs.5807	GTPase Rab14
	132883	AA047151	AA373314	Hs.5897	Homo sapiens mRNA; cDNA DKFZp586P1622 (f
	132968	N77151	AF234532	Hs.61638	myosin X
	132989	AA480074	AA480074	Hs.331328	hypothetical protein FLJ13213
60	132999	Y00787	Y00787	Hs.624	interleukin 8
	133071	T99789	BE384932	Hs.64313	ESTs, Weakly similar to AF257182 1 G-pro
	133076	W84341	AW946276	Hs.6441	Homo sapiens mRNA; cDNA DKFZp586J021 (fr
	133099	L09209	W16518	Hs.279518	amyloid beta (A4) precursor-like protein
65	133147	D12763	AA026533	Hs.66	interleukin 1 receptor-like 1
	133149	T16484	AA370045	Hs.6607	AXIN1 up-regulated
	133161	AA253193	AW021103	Hs.6631	hypothetical protein FLJ20373
	133200	AA432248	AB037715	Hs.183639	hypothetical protein FLJ10210
	133220	X82200	NM_006074Hs.318501		Homo sapiens mRNA full length insert cDN
70	133260	AA083572	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fls, clone R
	133295	L00352	AI147861	Hs.213289	low density lipoprotein receptor (famil
	133349	N75791	AW631255	Hs.8110	L-3-hydroxyacyl-Coenzyme A dehydrogenase
	133391	X57579	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a
	133398	X02612	NM_000499Hs.72912		cytochrome P450, subfamily I (aromatic c
	133436	H44631	BE294068	Hs.737	immediate early protein
75	133454	AA090257	BE547647	Hs.177781	hypothetical protein MGC5618
	133478	X83703	X83703	Hs.31432	cardiac ankyrin repeat protein
	133491	L40395	BE619053	Hs.170001	eukaryotic translation initiation factor

	133510	AA227913	AW880841	Hs.96908	p53-induced protein
	133517	X52947	NM_000165	Hs.74471	gap junction protein, alpha 1, 43kD (con
	133528	M11313	AU077051	Hs.74551	alpha-2-macroglobulin
5	133538	L14837	NM_003257	Hs.74614	tight junction protein 1 (zona occludens
	133562	M60721	M60721	Hs.74870	H2.0 (Drosophila)-like homeo box 1
	133584	D90209	D90209	Hs.181243	activating transcription factor 4 (tax-r
	133590	T67986	T70956	Hs.75106	clusterin (complement lysis inhibitor, S
	133617	AA148318	BE244334	Hs.75249	ADP-ribosylation factor-like 6 interacti
	133651	U97105	AI301740	Hs.173381	dihydropyrimidinase-like 2
10	133671	T25747	AW503116	Hs.301819	zinc finger protein 146
	133678	K02574	AW247252		nucleoside phosphorylase
	133681	D78577	AI352558		tyrosine 3-monooxygenase/tryptophan 5-mo
	133722	X53331	AW969976	Hs.279009	matrix Gla protein
	133730	S73591	BE242779	Hs.179526	upregulated by 1,25-dihydroxyvitamin D-3
15	133750	X85735	BE410769	Hs.75873	zyxin
	133802	L16862	AW239400	Hs.76297	G protein-coupled receptor kinase 6
	133825	U44975	BE816902	Hs.285313	core promoter element binding protein
	133838	M97798	BE222494	Hs.180919	inhibitor of DNA binding 2, dominant neg
20	133859	U86782	U86782	Hs.178761	26S proteasome-associated pad1 homolog
	133889	AA099391	U48959	Hs.211582	myosin, light polypeptide kinase
	133960	M19267	M19267	Hs.77899	tropomyosin 1 (alpha)
	133975	D29992	C18356	Hs.285944	tissue factor pathway inhibitor 2
	133977	L19314	AI125639	Hs.250686	halcy (Drosophila)-homolog
25	134039	S78569	NM_002280	Hs.78672	laminin, alpha 4
	134075	U28811	NM_012201	Hs.78979	Golgi apparatus protein 1
	134081	L77886	AL034349	Hs.79005	protein tyrosine phosphatase, receptor t
	134164	C14407	AW245540	Hs.79516	brain abundant, membrane attached signal
	134203	M60278	AA161219	Hs.799	diphtheria toxin receptor (heparin-bindi
	134238	R81509	AA102179	Hs.160726	Homo sapiens cDNA FLJ11680 fs, clone HE
30	134299	AA487558	AW580939	Hs.97199	complement component C1q receptor
	134332	D86962	D86962	Hs.81875	growth factor receptor-bound protein 10
	134339	AA478971	R70429	Hs.81888	disabled (Drosophila) homolog 2 (mitogen
	134343	D50683	D50683	Hs.82028	transforming growth factor, beta recepto
35	134381	U56637	AI557280	Hs.184270	capping protein (actin filament) muscle
	134403	M61199	AA334551		sperm specific antigen 2
	134416	M28882	X68264	Hs.211579	melanoma cell adhesion molecule
	134493	X15183	M30627	Hs.289088	heat shock 90kD protein 1, alpha
	134558	S53911	NM_001773	Hs.85289	CD34 antigen
40	134817	U20734	AU076592	Hs.198951	jun B proto-oncogene
	134893	D28235	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p
	134899	AA236324	AW968058	Hs.92381	nucleoside diphosphate linked mol
	135052	AA148923	AL136653	Hs.93675	decidual protein induced by progesterone
	135062	AA174183	AK000967	Hs.93872	KIAA1682 protein
45	135069	AA456311	AA876372	Hs.93961	Homo sapiens mRNA; cDNA DKFZp667D095 (fr
	135071	L08069	W27190	Hs.94	DnaJ (Hsp40) homolog, subfamily A, membe
	135073	AA452000	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f
	135170	AA282140	T53169	Hs.9587	Homo sapiens cDNA: FLJ22290 fs, clone H
	135196	J02854	C03577	Hs.9615	myosin regulatory light chain 2, smooth
	135348	AA442054	U80983	Hs.268177	phospholipase C, gamma 1 (formerly subty

TABLE 4A

Table 4A shows the accession numbers for those plays lacking unigeneID's for Table 4. The plays in Table 7 lacking unigeneID's are represented within Tables 1-6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

5			
10	Pkey: CAT number: Accession:	Unique Eos probeset identifier number Gene cluster number Genbank accession numbers	
15	Pkey	CAT Number	Accession
20	100752	33207_21	T81309 BE019033 R94181 BE019198 NM_000612 J03242 AW411299 BE300064 BE297544 R94182 AW630108 T53723 D58853 H78073 H80594 BE299560 T48899 H70196 M17426 N77077 S77035 H56384 H61664 H78540 T84527 C17188 H60255 H71980 R92644 W79050 X00910 M29645 R91055 M17863 M17862 T71815 BE299561 BE484561 X06260 R94741 T54218 C18594 BE262015 X06161 AW409889 AA378400 BE263228 BE313278 R88116 BE313457 H43500 T48617 BE313781 H77309 AJ207601 X06159 H40413 X03425 T87663 R10627 X03562 M14118 W03982 R97520 H81229 T83157 H83168 H48762 AA669898 BE263054 H47289 AA022807 R11555 H74260 R76968 R28338 H72534 H72464 H62031 N72478 H45355 AW411300 R89113 R69135 H58454 T83281 R93476 H69645 H68015 T82229 H71089 T85121 H59939 W65299 N78176 H53909 N72373 R21788 H04660 H59639 H61874 BE262219 T53614 T73335 N50464 W00943 N77189 R89257 AA570502 R89432 R06366 AA553480 AA776271 AA551359 AA551050 H51670 AA601052 BE299081 H68198 H52276 BE207832 N91192 H70332 X07868 X07868 H69464 H53782 H73710 R80435 AA553384 AW884176 N53475 T71662 AW954036 AW954033 AA552931 H93206 AA430218 AA553476 A918470 T54124 BE207982 BE300177 N73994 AW882625 N39549 N53838 AA722389 H71878 H58909 H37849 H78435 T47933 R77174 R83814 AA411890 H94199 AA663208 BE205778 AA490137 H70492 R98232 H37800 AA679294 H40341 H74238 H47290 H73231 T48618 AA025428 AI039521 H92969 N59389 H80538 H72933 T90630 AA411891 N55000 H74225 AA340290 AW957061 T54316 AA340437 H57125 H58908 H79027 H63450 N74623 R93425 H68714 H68758 H68396 H48763 H69256 H57320 W53831 H53589 N68833 N52453 H56048 H69870 H78074 R69253 R83375 T53615 H94330 H58455 H90864 T47934 H74261 R89258 R97997 R91056 R28339 R86760 H78235 R97521 H67692 H40358 AA022688 H52513 H59601 T86690 H65256 H63397 W65397 AA553588 R19280 N52645 W73930 R06367 R21743 H72372 N73921 AW883539 AW882639 T40616 H47084 R95723 AA634316 AA862781 H77310 R91389 H93111 R92767 T54512 R89341 H70333 H57817 H82941 H62032 N52638 H58385 T91796 H51086 AA340292 T49918 H81230 R36121 N50411 T87684 N62436 N39340 AA665637 AA340446 H93377 H92973 BE296290 BE269788 H61665 AA340444 N54605 AA454101 R10628 R94200 AI200549 AA342640 BE298855 BE250229 T49918 H82008 N28278 AW880662 H71268 N76791 H47685 H55255 W05198 AW889144 N76677 H71702 H68036 H71915 R91612 R87807 H68059 AI133328 AI247866 AA621443 AW881050 AA700847 AA340413 AW878608 AW881181 AW878249 H71918 N54596 BE161581 AW878082 W04212 AW881040 AW885492 AW880519 AA334887 AW878715 W06882 AW830222 AW885381 H70869 AW381778 H47601 AW889982 H63868 AW884986 AW878713 AW878685 R36391 AW878694 AA368070 C03393 AW878695 AW878705 AW878665 AW878742 AW878820 AW878823 AW878688 R29048 AW878690 AW878686 AW878810 AW878827 AW878733 AW878659 AW878749 AW878681 AW883353 AW883277 AW883300 AW883565 AW883298 AW883143 AW883045 AW883482 AW883352 AW883417 AW883357 AW883231 AW883474 AW883355 AW882620 AW882533 AW883754 AW883139 AW882827 AW883641 AW883567 AW883481 AW882983 AW882982 AW882465 AW883419 AW882466 AW883639 AW883230 AW882981 AW882534 AW882874 AW882619 AW883480 AW882826 AW882831 AW882835 AW882830 AW883563 AW882456 AW627642
30			W73853 AA928112 W77887 AW889237 AA148524 AT749182 AT754442 AI338392 AI253102 AI079403 AI370541 AI697341 H97538 AW188021 AI927669 W72716 AI051402 AI188071 AI335900 N21488 AW770478 W92522 AI691028 AI913512 AI144448 W73819 AA604358 N28900 W95221 AB68132 H98465 AA148793 AW980564 AA092457 T55690 D56120 T92525 AJ815987 BE182608 BE182595 AW080238 M90657 AA347236 AW981686 AW176446 AA304671 AW583735 T61714 AA316968 AI446615 AA343532 AA083489 AA488005 W52095 W39480 N57402 D82638 W25540 W52847 D82729 D58990 BE619182 AA315188 AA308636 AA112474 W76182 AA088544 H52265 AA301631 H80982 AA113786 BE620997 AW651691 AA343799 BE613669 BE547180 BE546656 F11933 AA376800 AW239185 AA376086 BE544387 BE619041 AA452515 AA001806 AA190873 AA180483 AA159546 F00242 AI940609 AI940602 AI189753 T97663 T66110 AW062896 AW062910 AW062902 AI051622 AI828930 AA102452 AI685095 AI819390 AA557597 AA383220 AI804422 AI633575 AW338147 AW603423 AW606800 AW750567 AW510672 AI250777 AA083510 AW629109 AW513200 AA821353 AI677934 AI148698 AI955858 AA173825 AA453027 AI027865 AW375542 AA454099 AA733014 AI591384 R79300 R80023 AA843108 AA626058 AA844898 AW375550 AA889018 AI474275 AW205937 AI052270 AW388117 AW388111 AA699452 AI242230 N47476 H38178 AA366621 AA113196 AA130023 H39740 T61629 AI885973 AW083671 AA179730 AA305757 AI285455 N83956 AA216013 AA336155 AW999959 T97525 AA345349 T91762 AA771881 AI285092 AI591386 BE392486 BE385852 AA682601 AI682884 AA345840 T85477 AA292949 AA932079 AA098791 D82607 T48574 AW752038 C06300 R20840 R20839
35			BE273749 BE397561 BE387189 AL037858 AL037878 AI963094 BE259216 AA011363 AL036189 BE562325 AA251169 BE617431 N98537 AA158093 AL047800 M34539 NM_000801 AA312140 D16971 AA158904 AA307114 AA312803 T09203 AW629686 AL048504 BE388578 AA220957 AA158364 BE267385 AA294971 C18055 BE241757 AA115056 AI936769 BE378435 BE206971 AW874924 BE622060 AA604674 AA115273 AW402159 AA338608 BE588819 M80199 X55741 AA375111 AA376016 BE612671 AA805742 AW405588 N25850 N44580 H06031 AW403549 BE536552 AA056726 H543239 AA082517 AI201645 AI201842 AI192622 N40104 AA370921 BE547569 AI969602 AA302038 AI197890 AW268354 AI014938 W54448 AI541395 AA037272 BE538826 AL039613 BE536130 AA299355 AW805147 AW974624 H53220 AI471471 AA399303 AA007386 W35106 BE613277 R12739 R12738 AA304342 AA687802 BE409581 AW98844 AV682092 AW904105 AA011375 BE315214 H99302 BE537893 N32299 AW855829 AI291320 BE078322 AI301395 AA303362 N32719 AA358328 AA357877 AI952540 H56279 H02758 H02048 AW805233 R82224 AA410772 AA291352
40			
45			
50	117156	145392_1	
55	131859	3672_1	
60			
65	125565	1704098_1	
70	133607	1227_6	
75			

			BE171109 N69935 BE169248 AA361173 H44978 BE817887 D62560 AA084043 W03595 R67219 N36477 M42924 R67104 H44901 H79695 W21105 AA393988 W30899 AA316096 BE522896 W46872 AA442678 BE544893 BE540112 BE521873 AA338067 N55062 BE398154 BE621210 AA740760 C03739 C03206 BE396692 AA482370 AA031614 AA301575 AA304710 AA132153 AA029796 AA994960 H19567 AA42969 H49781 H46871 AA035395 AA056185 AA149378 5 AA643080 AL135479 AA292329 AA654337 AA041228 AA454888 AA025039 W58331 AA625981 T94941 AA302448 H19900 AA218956 AA513790 AA563962 AA398076 W44441 AA293278 W47373 AA625879 W30688 AA043029 T64284 R79151 AA304340 AA485186 AA604939 R82470 AA421425 AW771456 AI339329 AA304424 AA606236 AA936934 AA587673 AI209162 AI697301 AI479995 AI679814 AI361950 AW189125 AI955888 AI986019 BE301019 AI084792 10 AI310211 AW189307 AI022070 AW977204 AI146825 AW190163 AW303281 AI828345 BE046043 AW029257 AA482268 AI246507 AW20729 AW084932 AW439514 AI890487 AW436992 AI523896 AI186612 AI659953 AI889773 AA687527 AW072694 AW262153 AW467371 AI613269 AI679238 D54404 AA158103 AW105527 AW149739 AW150361 AW268387 AW117708 AI951682 AI687440 AW674285 AA878365 AI587082 AA732095 AA018899 W45681 AA627300 BE113304 AA765691 AA612935 AI814658 AW316916 R66594 AA514640 AA025040 AA031472 AW732076 AA029797 AI244560 AI128734 AW381720 AI023260 AI263283 AW613175 AI890675 AI720156 AW631348 AI635106 AI278045 AA303979 15 AA703505 W45449 AW078661 AI292052 AW381707 AI147854 AW381743 AA158905 AA303258 AA888144 AW195967 AA428706 AA989559 AA617731 H19882 BE543418 AA830386 AA421302 W58652 T94995 AI869743 AI679145 AW085971 N98425 AA765136 AI347027 AI356955 AA928038 AI679717 AA458459 AA679281 AI367973 AI270041 AA765135 AA732793 AI798447 AA668646 AA251008 AA845338 AI401737 AA056186 BE043308 AW662375 AI302110 NS0724 W96332 BE537047 N26983 AI567172 AA765296 AW673237 N28784 AA534275 AA084044 AW087973 20 AW300766 T63398 W46823 R39790 AI364185 AW298582 AA454814 AW069878 N67751 H05982 N23140 AI362647 AI302086 AI677772 N25755 H53114 AA706133 T93511 AA429291 AA935294 AA987647 W02803 R66595 AI808795 W23673 W444079 AA722672 H49538 AW131042 AA531603 AA908665 AA040791 AA235312 W52205 N93444 R82180 H02759 H79698 AW088894 H56079 AA961143 AW067776 AW973745 AA016311 AW071227 AA017511 AI753994 W47374 T64155 AA296092 AI698626 AA558158 AA296088 AW794259 H01983 AA149267 AA485076 AA975856 H44938 25 AA035396 AI955555 H46289 AA488161 AI631222 AA359047 AW794253 AI808962 AW243930 AA526145 AW878734 AA018464 AA132031 R67220 R79152 AA296093 H54300 AI005160 BE242548 AW992803 AW878644 AW878666 T27742 R82471 AW517604 AW472738 AI282904 R39791 AA486098 AW467891 AW960520 AA551738 AA056821 AW945197 R66373 AA554236 BE242202 AI904376 AI632590 H19484 R00890 AI627677 AA302287 AI869451 AI734855 AI708073 AI832902 AA585184 AW204299 AA055565 D12417 D11975 T63543 AW684099 R54423 BE612712 T96340 T63985 30 AA598917 T40735 T64053 AA149284 AW272548 AA363445 AA042893 AW300697 BE261973 T35301 T53500 AW878729 AW878657 AW794391 AA069193 R01553 H44875 AA385406 AA533968 M93060 AL135600 W96331 AA017651 AA018849 AA017692 H53337 BE278690 AA731598 AA018512 AI076813 AI022644 R02585 X52220 AW296894 AA825671 AI699321 AI939601 AW592611 AI146747 AA608921 AA158365 AW590007 AA354519 D20081 R02704 AW798339 M92422 AA094903 AA007676 35 133681 13893_1 AI352558 Z82248 X78138 NM_003405 AU077248 AA223125 S80794 D78577 AI124697 AW403970 BE614089 BE296713 BE621334 L20422 X80536 D54224 D54950 X57345 N29226 AA127798 AA304253 F08031 AA192540 H67638 AA321827 AW950283 AA084159 BE538808 AW401377 AA256774 C03366 W46595 W47608 AA305009 H69431 H69456 AL120082 H11706 AA303717 AA361357 H22042 H78020 AW999584 AA134368 AA322911 AA322961 H60980 N85248 N31547 H79624 T11718 W58286 AW894663 AW894624 BE187441 BE170015 AA304626 AW602163 AW998929 AA156681 40 AA151067 BE002724 AA608688 H82692 BE155392 AW383636 BE155394 AA487004 AW383504 AI342365 R82553 W16498 BE155344 AI143398 R69901 AA322873 AW3040648 R25364 AA367935 AI559406 AA033522 AA374252 AW835019 AI922133 AI697089 N89662 AW189078 AI199076 AW151598 W59944 AA662875 W94022 AA299055 AI039008 BE29449 AA583503 AI635674 AW131665 AA73820 AW273118 AW900930 AA908944 AI688035 AW170272 AI082545 AW468176 AI0808761 AI082748 AI911682 AI248943 AI831016 AA192465 AI218477 AA938406 AA385288 45 AI809817 AA905196 AI791245 AI470204 AI188296 AI421367 AI125315 AI087141 AA629032 AA740589 AI554181 AA150830 AI248541 AI077943 AA775958 AA864930 AI261478 AI123121 AI310394 AA862331 AA872478 BE337084 AI205606 AA720684 AI872093 AW150042 AL120538 AA219627 AA988608 C21397 AI359337 H25337 AI089749 AA605146 AI359620 AA150478 AI359738 AW383642 AW995424 I766457 R56892 AI089839 W61343 M69107 W46459 50 AA655955 N20527 AI279782 W46596 AA776573 H23204 AI866231 AI083995 N21530 AA126874 D82630 W65437 AI086917 AW382095 AI086877 H69844 AW340217 W85827 L08439 AA262704 AA505380 W47413 W94135 AA223241 AW089153 AA084101 BE338000 AA096126 T28031 AA491574 R84813 AA774536 AW383522 AA155615 AW383529 AA491520 AW028427 AA171496 AI469689 AW664539 AI811102 AI811116 BE464590 BE350791 H78021 T15405 H21979 AA219489 H13301 AA505883 AI864305 AA23963 AW084401 F04963 R69858 H67097 AI917740 AI655561 H69864 55 AA033631 AW383484 AI886261 H25293 AA513281 AW271187 H11617 N79982 AI174338 AI904207 AI904208 BE614558 W94127 W65436 AI272249 AA700018 AI579932 AI085941 AW152629 AA334551 BE008229 AA307537 AW961158 AW995894 AW995826 NM_006751 M61199 AA045603 AL036372 AV645606 AI688095 AW351901 AA101337 AA101345 N73342 BE018030 BE569044 AW841975 AA373388 BE090412 H95440 60 N53845 R67967 AA093441 AA363427 H93708 AW023134 AW994988 AW994989 BE090429 R23614 AI567932 H03726 H01101 H01867 AA548743 AI671806 AW872949 AW872941 AA742447 AI199788 AA045604 AI637465 AI741796 AW242217 AW131463 AI765302 AI683923 AA889762 AI804889 AI988437 C06049 BE502340 AI695651 AI491970 AA496804 AA281008 AA665699 AI473814 BE301445 AA707837 AA551925 AI017348 AI208185 AA775203 AA156286 AA557463 H95441 AA768547 AW769358 AA991197 AA181954 AI091389 AI147289 AW771837 AI638582 AA844411 65 AI374750 T28320 AW951272 AW085923 H02834 AA843259 AA814696 AW183290 AA158453 N68125 N69039 AA100423 AA101346 AI918720 H01102 R67868 H01868 N66438 R46580 AI858433 AA599560 AA187577 AA157481 AA361520 AL047827 AA158452 R21688 AW964874 AA325161 R40871 AW752395 AW375924 R13355 AA281174 AA428908 AW450979 AA136653 AA136656 AW419381 AA984358 AA492073 BE168945 AA809054 AW238038 BE011212 BE011359 70 126872 142696_1 BE011367 BE011368 BE011362 BE011215 BE011365 BE011363 AA404418 AI217248 121335 279548_1 AA353093 AW957317 AW872498 AI560785 AI289110 AW135512 X97261 T68873 130018 18986_1 AI743860 N49543 AW027759 BE349467 AI656284 BE463975 R35022 AA370031 AW955302 AL042109 N53092 AI611424 121822 244391_1 AL079362 AI969290 AI928016 BE394912 BE504220 BE467505 AI611611 AI611407 AI611452 W56347 AI284566 AI583349 AW183058 AI308065 AI074952 AA437315 AA628161 AW301728 AI150224 AA400137 AA437279 AI223355 AA639462 AI261373 AI432414 AI984994 AI539335 AA401550 AA358757 AI609976 AA442357 AA359393 AA437046 AA370301 AA428328 AW272055 AI580502 AI832944 AI038530 AA425107 AI014986 AI148349 AW237721 AW777956 75 123523 genbank_AA608588 AA608588 AW137677 AI125293 AA400404 R28554
--	--	--	--

123533	genbank_AA608751	AA608751
125091	genbank_T91518	T91518
123964	genbank_C13961	C13961
102491	entrez_U51010	U51010
5	118475	genbank_N66845
118501	genbank_N66905	N66905
113947	genbank_W84768	W84768
101447	entrez_M21305	M21305
101667	13349_1	NM_005381 M60858 AW373732 AW373724 AW373689 AW373629 AW373609 AW373776 AA187806 AW386946
10		AW374207 T05235 AA216203 AW385556 AA308940 AA308526 AA315461 AL036757 AW373711 AW403124 AW403640
		AW377084 T27360 H62638 F06957 AW377051 AA554779 AA378568 AA096007 AW352407 AW302637 F07929 H17433
		AW382712 H06665 F07292 N39875 AA089729 H62558 N42842 R12952 AW373735 AW364155 AA056183 W39185
		AW382708 N32488 AF114096 AW375993 A1135569 W52561 AA603040 AA133710 A1928796 AW176370 AA827519
15		AW338437 AA521142 T29341 A1800461 AW317002 AA703914 AA660830 A1859203 A1445772 AA714334 A1817066
		A1832027 AW510442 A1635802 AW088306 AW088672 AW408555 AW467542 AA552657 AA152387 W32081 AA582124
		AA074040 AA931657 A1051154 AW410203 A1921644 H17434 A1832330 AW404836 A1925038 AA088423 AA954166
		AA580453 AW021292 A267215 AW080082 AW383778 A1933053 A1919097 W31557 N90245 AA931591 AA563995
		F36352 AA056184 AA476294 AA641327 AA533550 A1749630 W58323 AA569119 AA508573 A1809050 A1378996
20		AA411362 AW407505 AA938104 AA074041 AA632876 AW193748 AA507873 A1270128 A1472365 AA411363 A1523216
		A1719965 A1816302 AA182681 A1707990 AA133588 A1758537 W60253 AA60308 AA135423 A1083904 F04188 N86963
		AW408778 A1678595 A1270568 AA722059 W58234 F33650 AA090547 AA285108 AA425981 N85079 D20218 A1273980
		AA159028 F03228 AW247914 N26918 AW272741 N90109 H05666 N23327 AW247953 R44748 AA962015 F03558
		A1752394 AW409913 AW248398 A1816463 A1752393 AA325370 AA263089 A1570130 A1971951 A1160658 A1357360
25		AW168686 AL121075 AW050536 N21672 W67748 AA514242 A1127386 H14607 A1185752 W79364 AA088520 AA152476
		AW351940 AW373683 A1940524 AW374953 T56500 N24329 A1940720 AW374933 AW374947 AW391913 AL138337
		AW376241 AW062943 F26666 AW410202 AW062958 F34529 AW381807 AW393315 W17147 AW176359 AA664576
		AW380424 AA306040 A1745674 AW300951 A1188579 A1438973 A1305271 AA433818 AA612807 A1831809 A1940409
		AA158663 A1572988
30	108931	genbank_AA147186
	103138	entrez_X65965
	103432	entrez_X97748
	119174	genbank_R71234
	133678	11235_1
35		AW247252 AA346143 NM_000270 AA381085 N91995 X00737 AA381079 AA296473 AA296110 AA315735 AA311617
		AA328750 AA376804 AW403290 T95231 M13953 T47963 H82039 AA279899 AA627997 N76320 N89527 H37842
		W20095 AA457308 AW469547 AA724143 H83220 AA319496 W86334 W30892 R89169 R99427 N41854 H47286
		AA348094 AA045089 R63016 A1922219 A1024906 A1096488 A1885005 AA194872 N90489 A1452544 H72411 AA282427
		AA430735 R68963 R22453 H70385 AW129369 AW467320 AW519082 AA345018 AA582183 A1961789 R65918 N30611
		A1979189 A1280889 AW273191 R66531 A1285845 A1675927 A1421990 AW190879 H37794 AA699667 H68427 AA954388
40		A1188757 A1140048 AA430382 A1204151 AW247864 AA559099 A1431420 AA548276 A1149466 AA72669 AA694388
		AA724168 AA301651 AA281952 AA779925 AA234760 W86290 AA913603 AW511745 A1500697 AA814922 AA835040
		T47864 H53998 AA975804 R98710 A1077604 N70252 R98084 AW250171 H69268 A1597614 AA970746 AA972548
		A1377116 R62962 H16737 R89070 AA731329 R66532 N54354 A1818832 H81944 N71567 T85122 W86463 AA437095
		A1431999 A1915724 N63851 A1674743 AA457307 AA211475 N64444 A1799146 H72853 R99335 H60413 AA770367
45		AA156105 A1269937 H64029 H89728 R65819 AW470496 A1873318 A1735713 H82987 C02447 A1478666 T27651
		A1699770 AW025156 H69719 A1984717 N69225 A1459856 AA953577 A1424691 H13843 R22404 A1873796 A1336002
		N70898 A1420854 AA541792 AA346142 A1000814 A1828348 AA045090 T51257 N90434 H13890 N73184 A1708083
		AA781606 AA329050 AA339985 R68964 H64795 W04186 H16845
50	119416	genbank_T97186
	119559	NOT_FOUND_entrez_W38197
	123473	genbank_AA599143

TABLE 5:

5	Pkey:	Unique Eos probeset identifier number			
	Accession:	Accession number used for previous patent filings			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
10					
	Pkey	Accession	ExAccn	UniGene	UnigeneTitle
15	115819	AA426573	AA486620	Hs.41135	AA486620
	132837	D58024	AA370362	Hs.57958	AA370362
	101545	M31210	BE246154	Hs.154210	BE246154
	102898	X06256	NM_002205	Hs.149609	NM_002205
	101192	L20859	BE247295	Hs.78452	BE247295
	102915	X07820	X07820	Hs.2258	X07820
20	105330	AA234743	AW338625	Hs.22120	AW338625
	107385	U97519	NM_005397	Hs.16426	NM_005397
	102024	U03877	AA301867	Hs.76224	AA301867
	134416	M28882	X68264	Hs.211579	X68264
	103036	X54925	M13509	Hs.83169	M13509
25	104865	AA045136	T79340	Hs.22575	T79340
	106124	AA423987	H93366	Hs.7567	H93366
	105330	AA234743	AW338625	Hs.22120	AW338625
	109001	AA156125	AJ056548	Hs.72116	AJ056548
	104764	AA025351	AJ039243	Hs.278585	AJ039243
30	133200	AA432248	AB037715	Hs.183639	AB037715
	105263	AA227926	AW388633	Hs.6682	AW388633
	105178	AA187490	AA313825	Hs.21941	AA313825
	109456	AA232645	AW956580	Hs.42699	AW956580

TABLE 5A

Table 5A shows the accession numbers for those pkeys lacking unigenelD's for Table 5. The pkeys in Table 7 lacking unigenelD's are represented within Tables 1-6A. For each probe set we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.		
Pkey:	Unique Eos probe set identifier number	
CAT number:	Gene cluster number	
Accession:	Genbank accession numbers	
Pkey	CAT Number	Accession
115819	10241_1	AA486620 AF205940 AA297524 AB034695 AA081335 NM_016242 AA188323 AA297537 H88204 AW953081 W31695 AW582203 AA248250 AW681211 AA426230 AA464807 AA426155 N44141 AA347390 AA770661 AI333225 N36136 AW665724 AA431894 AJ374976 AJ400254 AI338446 AA186695 H88205 W04527 AA487066 AW051414 AA918363 AA426573 AA425620 AW438654 AA090513 BE167284 BE167291 AI301726
102024	14505_1	AA301867 AW957981 R27614 AA155808 AI820990 AJ740711 AA301026 AA301015 AI220981 AI857670 AI537140 AW015210 AA030000 W46890 H44021 AI355967 AJ551735 AA058479 AA146932 T58265 R58590 AA047810 AA017387 AW026093 AA971133 AI827263 AI056416 AI355994 AI127691 H46603 U03877 NM_004105 AA157357 H42844 AA146824 AA187709 AA187269 AA304348 AA147292 AA361687 AA156041 AA330636 R32929 AA321130 AW950260 AA082157 AA029129 AA303708 AA028155 D31561 T84689 AA302493 BE153057 BE153181 W39408 AA187200 BE153250 AW383337 AW382622 AW382647 AW750072 BE153060 AW382630 AW371865 AW392464 AW382664 AW382658 AW382650 H61647 AW365075 AW365049 AA373397 BE072779 BE072781 Z30254 W24381 BE153254 AA044442 BE072729 BE072731 N94740 AA146945 AW802737 AI826799 AI085395 R34034 H65140 AA082800 H88275 AA147824 R63882 W80899 AA296413 AI765300 AI862426 AW022055 AW300003 AI743784 AI862635 AI985428 AA147764 AW573245 AW190290 AI040898 D57613 N63457 AA148082 AI028458 AA148110 AW814489 N75105 AW629443 AA704122 AW582220 AA181240 AA057495 AA182224 AI261751 AW388595 AI472205 AW470672 AA102546 AA789046 AA182416 AA062668 AW300732 AI288220 AA181982 AA146825 AA028130 AI985522 AA303344 AA081313 N69082 AA182035 AI867128 AA100902 AA605087 N67178 AW020324 AW890446 AI472191 AI335691 AI597837 AI081143 AI335681 AA040443 AI128067 AI878244 AA018303 AA157260 W80792 AI934590 AI096430 T54343 AI446350 AA165196 AA780683 AA603631 AA047787 AA968580 AA912645 AW890504 AW026913 D56983 H52088 AA156121 R30848 AW023036 AI590980 N67345 AI753225 AI753283 AI183768 AA147818 H89101 AI362141 H89205 AI147711 AA321129 AA668622 AA343479 AW069438 AI422376 AW629270 AA013413 AI221948 AA970605 N52335 H38366 T91180 AA657841 AA017386 AA152227 AA187593 AI913340 AI719313 AI969943 AI701271 AJ004328 AI868348 N93659 H65093 H25736 D57007 D56957 C00987 D61839 D56661 AI472137 AI971002 D56971 BE048830 D57972 AI589286 AI361055 AI361071 AI292223 AA155898 D57139 D57981 D57345 AI420034 D57332 D57959 AA875933 R33493 N67558 D58353 AA188394 AA147966 AI160640 AI363165 H40638 AA578137 AW950265 AA300943 AI128899 H46584 AA917355 N57820 AA320504 H51959 H25737
101545	24607_1	BE246154 M31210 NM_001400 AA193392 NM_016537 AF233365 AF022137 H27787 AA370448 F05373 T27666 W21494 AA036907 AI249966 N93476 F01623 AA304390 AA308808
109456	180833_1	AW956580 AA886361 AI147670 AI090115 AI166863 AA232645 H99504 AA374707 AA380875 AW139567 AI735132 BE439385 AW629780 N28322 AA232789 AA232790 N73285
103036	17145_1	M13509 X54925 NM_002421 M16567 X05231 M15996 W39354 AA186634 AA852324 AA187507 AA081149 AA186524 AA187264 AA187361 AA386155 AA186973 AA374217 U78045 AA081230 AA188049 AA186393 W56827 AA852602 AA157468 AA308204 AA186754 AA186808 AA082516 AA304334 AW376428 BE439384 AW376420 AA156273 T18504 AA186521 W49496 AW084608 AA083575 AA372360 AW963590 AA132297 W47445 AA186376 AA157628 AW003999 AI037890 AI858060 AI589010 AI743739 AI452673 AW304188 AW117854 BE439933 AA157416 AW778966 AI038497 AA081006 AA100829 AA181048 C02231 T27821 W23960 AW954802 AI471432 AW801296 AW801289 AW801603 AW801523 AW801292 AW801542 AW801601 AA181134 AI445147 AA191501 AA582862 N94407 AI147810 AA181880 W49497 W52714 AA188249 AI832881 AI082493 AA503656 AA182682 AW801393 AA182830 AA181882 AA182826 AI613182 N94510 W47343 AI085755 AI076956 AI918428 AA081208 AI282835 AA147528 AI081490 AI654536 AA181875 AA081282 AA186389 C06085 AA083542 AI800644 AA157642 AA101069 AA157752 AA158121 AA143331 AA081283 AA852603 AA188296 AI932880 AW449628 AA187348 C02091 AA514656 AA082736 AA308786 AA143201 M16567 AB037715 AI351347 AI375796 AI884765 AL121124 W01068 AI807275 T85240 R42807 AW515645 AI057314 AI033520 AA057671 N70215 AA054215 AW204183 AA552149 T95130 AW796310 AI866520 AW275564 AW796308 AI637901 AW197404 T78408 AA456232 AW206463 AA779800 AI052696 AA026744 AA454623 AW470729 R45490 AW770258 AI038393 AI290170 AA722734 AL121125 R41608 AI862414 AA838611 R45582 AI278083 BE468849 BE219944 AA418030 BE041555 AA578572 T16528 AW006344 Z39782 AI244848 AW137344 AA707400 AI032028 BE540464 AI094265 AI184281 AA931890 AW382744 AW382729 AW020448 AW827237 AA431226 AI672059 AW772345 N70172 AW022003 AI862704 H19344 R61511 AI080204 H16566 AA432248 AI767980 T16688 AI984342 AI217478 AI767095 Z38551 AI359568 AI361437 AI041000 R07033 H16608 H19054 R12874 R61567 N98368 BE221199 Z42320 AA094554 R07078 AW860885 AA418090 R41262
132837	256666_1	AA370362 AA364110 AW959554 AW371737 AW382068 AW604716 AW604713 AA487827 AW371674 AA429137 BE503321 T93570 W72803 AI093078 AA487977 AI241562 BE439445 AW204065 R51635 AI802994 T10362 W68553 AI866215 AW152154 AA700718 AI127443 R15824 AI537587 AW953110 D58024 AI520811 AA693670 AI453280 W76329 AW023955 AW022563
102898	24023_1	NM_002205 X06256 M13918 BE070866 AW239485 AW996127 BE273894 BE272590 BE410252 R25975 T11786 T11787 AA301142 AA301165 AW960506 BE272819 AA386086 T39391 AA285303 AA370580 D58585 W58668 AA156213 W24142 AA343323 AW796067 AA151197 AA376121 R94782 AA302363 H90357 R82621 AA301877 H55997 AW796059 W92358 AL046458 AA471198 AA301952 R46287 R82694 H03186 AA187708 R32562 R27094 R25847 R25320 AW949809 H13505 H79049 R32403 H11213 R39710 H49765 H21142 H21006 AA417664 W52075 N56771 AA284240 N98556 N30907

			AA707335 AW603761 A1340367 A1814584 AA524182 AA370076 AA418785 AA704082 A1806851 H25513 T56388 AA419627 H03986 H20963 T56245 A1459715 AW973768 A1334098 A1693020 T63414 R26246 AW167251 H55998 A1274916 AA778387 A1755253 A1033667 AW083222 AA181979 R26865 AA661627 AA706329 A1798648 AA612799 A160180 A1274973 A1039284 AA301880 A1042429 AA307632 A1085688 A1278366 A1498890 AA303865 A1954844 AA502380 AA156334 AA723480 A1803584 A1581026 AA304584 N51038 R94702 R69814 AW150962 A1570049 AA588807 AA151188 T53400 A1567709 A1185326 AA309205 AW338969 R53903 AA981891 AA301643 A1493337 A1026049 H25614 A1741075 R28632 AW168445 A1333068 H49978 H91267 AA558193 AW079663 AA627380 AA807401 A1199956 AA686118 A1718216 AW193228 A1077745 A1500496 A1266059 AW080383 R05468 R26757 R32404 AA716599 W92322 A1077734 A1270181 R46198 A1217540 AA304045 AA305421 AW074445 A1468256 AW089568 AW571605 BE162930 H41009 AA578313 AW874497 AA181284 AA861947 T29451 D20841 T58618 AA418731 A1282500 AW081407 AA604560 AA729855 A1262538 A1580225 X07820 NM_002425 BE271570 A1263526 AW296143 A1829878 A1873162 A1085155 AA857496 AA709305 C02220 X68264 NM_006500 AF089868 BE257461 BE275425 AW997154 A1902799 A1902803 M78206 AA085691 AW392972 AA325490 BE006161 AA349269 AA323568 AL042548 AA191148 AA187703 AA322791 A1297452 T11625 AW366487 AA303513 AA186961 AA173480 N28330 N28379 W40320 AA187118 H03695 AA402709 BE407476 H00354 BE276589 AA351284 AW379921 AL138060 BE410587 AA113094 AA340481 BE277483 R21191 R79518 N86170 AA320505 AA296065 AW951900 AA658897 AA650052 AA654304 AA191691 N26649 AW080963 A1265800 N72019 A1453458 AA092563 AA402310 A149450 A1661054 AA302358 T71566 AA302047 AA303432 N21289 H27357 AA303504 A1174583 AW151762 AA181958 AW880618 AA630773 A1888539 AW901058 A1373405 AA341941 AA086217 A1675590 A1053936 AA633570 AA897619 A1270656 N93847 N40689 AW517517 N20030 W95985 AA303955 H89170 AA309917 N21642 AA373132 W38517 A1687806 W76182 AA101065 AA036916 N45635 A1744510 A1669803 A1039157 A1126355 AA634607 AW131120 AW196638 AA190601 AA911130 BE221320 N92355 AA036752 H03696 AA588873 AA58886 A1041818 AA090477 A1093248 AA304755 AL137942 AL044688 A1083709 A1150965 N88891 AA635675 AA594898 W94657 AA182823 AW166205 F27886 R79246 F37329 AA565697 A1075739 A1088654 A1094287 A1204256 AA095203 T30320 AA688298 AA057324 N23442 AA075411 AA305046 A1031688 A1191503 AA111887 AA112264 N27929 AA187509 A1375522 A1474006 H06297 A1826177 N46880 H28333 AA075490 R22809 W79542 A1055934 AA042901 AA173481 AA301986 W47531 A1051747 AA187715 A1888888 AA93017 A1057530 T92954 N80227 AW273595 A1351260 AW170643 AW292979 AA302605 AA302330 BE349495 AA328602 AA302361 AA70984 AA155943 AA155914 AA313825 AW960347 AF223468 NM_016613 A1186345 AA186508 AA081195 AA147972 AA346943 AW961667 AA187722 AA187207 AW371052 AW449751 AW748803 AW391606 AW371047 AW371057 AW371085 AW362895 AW371092 AW377556 BE010930 A1016882 AA247878 C04398 C05158 F11398 AA188315 H23385 R55086 H15346 AA029106 AA228114 H17005 F08498 Z43376 AA095582 AA055186 AA463361 R15218 AA299132 AW103578 W21538 AA428131 AA187115 AA157197 AW371371 AA363562 AW965995 N56663 Z17878 AA228023 A1140342 A1100927 AA496988 AA055917 A1089303 AW014967 AW090248 AW338371 AW131066 D62983 D79713 A1583950 A136781 A1500705 A1471485 AW090239 D78784 D61847 D62789 D61842 A1086327 A1273381 D61815 D63043 A1913548 A1280560 A1510828 AA029996 C16343 C16513 A1075741 AW516308 A1804764 AA948068 A1356588 AW103452 AW573063 Z39445 C16489 A1949870 F04712 AA147623 AW026284 A1151538 AA081303 AA613890 A1251865 AW086499 AA992111 A1862091 A1373465 BE502094 A1922270 AA884288 AA157079 N56963 AW189145 AA428080 R55056 AA884068 AW771716 AA186662 C16364 H15723 A1921181 AA156888 H17006 AA187490 A1400994 AA346942 H28533 AW129047 R41656 H14636 AA995041 D58370 Z21131 D58186 A1383271 AA643977 D58044 A1934302 AW779425 F09065 H14930 AA890693 H23274 AW388633 AW378440 AW388283 AW388339 AW388333 AW388414 AW388413 AW388607 AW388453 AW388687 AW388480 AW388591 AW388711 AW388511 AW388438 AW388570 AW388449 A1694383 AW271455 A1652991 A1964041 AW366319 AW366321 AW961938 AW469211 A1634155 AA492188 A1624430 A1677965 N26502 A1933871 AW378431 AW378421 A1015391 AW352126 N59336 A1352317 AW197113 N67998 AW778935 A1476054 A1206826 R37116 R40211 AA227926 AA639698 R38073 A1001745 Z32854 A1619849 A1423703 F10774 AW388615 T16595 H05894 AW338626 R43226 R51640 A1307645 A1308100 A1085787 A1420357 A1692810 AA877160 A1953366 AA234743 A1039243 R68234 AA025351 AA971063 A1537757 AA025362 R81636 T86650 T79340 A1742317 AW182676 AW451460 AA20964 R43284 AA088179 AW590886 AW269529 AA045187 A1521736 A1827455 AA045136 AW271709 A1004344 AA639631 AA744417 AA744218 AA045357 AA045351 H93366 A1653547 AA336265 AW966175 BE566451 R71178 A1630656 AA234331 N55039 AA305632 AW960431 R34044 R32254 AW020970 AW451281 AW275041 A1636933 A1655640 AA423986 AA642466 A1684063 A1633876 A1624897 AA814795 AW590326 A1889166 AW243541 A1439691 AW473445 A1475516 AA741228 A1127534 AA165143 A1074714 A1654076 AA400674 A1560249 N50709 AW438621 A1806810 A1434579 A1308184 AA423987 A1141272 A1565586 A1338440 AA219628 A146643 A1885809 AA724260 AA633988 A1364172 A1798439 A1650801 R33503 A1435891 AA903649 T96161 AA665538 AA219620 A1309962 AA400707 BE247066 R32178 A1275982 AA661602 AW003197 BE466649 AA831198 A1620052 A1825387 A1634037 A1870978 A1670979 A1855092 R32304 AA828858 A1382428 AW023660 AA262892 T26891 AW089917 T26926 R32227 NM_005397 U97519 AW899329 A1902387 AA077792 AA078525 AW376607 AA077946 AA070415 BE208721 AW167958 BE293050 BE208240 A1646698 AA101314 BE393348 BE395122 AA077591 BE274036 AA313687 BE392220 BE378954 AA171461 AA464821 AW938242 AW938224 AW938243 AW938232 AA147953 N64294 AA205218 AW305065 AW517478 AA307983 AA377023 BE563629 R99976 N80294 T87719 T87928 AA496849 AA486344 AA204938 AW370448 AA318242 AW964384 H92423 W95317 BE378774 BE391156 AA349138 AA173095 AW513198 AA037672 AA148029 AA169726 W04791 AA075508 BE382937 BE395034 AF139793 AA961734 N48612 H64714 AW151251 A1565113 A1566881 AW087370 AA631168 AA622014 AW513098 A1857810 AW152287 A1052596 A1983246 AA024856 A1812456 A1677938 AW026403 AA972537 A1088497 AW999869 W094582 A140166 A1160659 A1566868 AA101283 AW190390 AW166466 A1401207 A1418156 A1625265 A146258 AW008592 BE223020 N58926 A1306797 AA037673 A1935992 A1304706 AA024939 A1216589 A1610423 A1354621 A1500677 A1679389 A1799310 N64508 A1128756 A1679897 AW589535 AA989333 A1500527 AA565479 AA913529 A1923295 F21691 AA989376 A1699064 AA902447 A1609010 AA772659 AA204983 A1337895 R99975 H65205 AA340766 A1339441 A1913855 AA450293 AW192010 AA070416 N72401 A1371481 A1247108 A1371261 A1364987 A1280171 A1269104 A1868756 AA909836 AA983640 A1973271 AA913092 A1868205 A1144112 A1190975 N58085 A1566638 N93405 AW150504 AW296846 A1687036 AA902984 A1824460 A1625047 AA653148 A1611228 AW131922 AA862687 AA902519 C01732 AW796045 AL044660 BE247295 AW068092 AL041313 AA159244 NM_005415 L20859 AL135570 W47073 AW516908 BE388271 BE408629 W46972 BE293646 BE256647 A1075010 AL041095 AA285300 AL039560 AA368740 W26602 AA399344 AA039235 W27631 AW834898 AW834914 R93390 AA378039 AV649660 T03674 N98824 AA399974 AW843378 AA368267 R08256
--	--	--	--

WO 02/079492

PCT/US02/04915

5

10

109001

146370_3

AV653575 RZ7900 N48215 AW366371 N45500 AV652967 AI889251 AI080457 N39021 AI738542 AW242849 AI857471
AI859775 AI582830 R75850 N66564 AW341636 AI499006 AI887217 AW026694 AW162840 AA039313 AA831346
AI393465 AW089210 AI743830 AA744243 AA401310 AW439758 AW088152 R93391 AA291379 AA225220 AW009358
AI192879 AA291202 AI565089 AA225089 AA807688 AI052058 AI341641 AI066625 AA333864 AA159147 AI923912
R75851 AI761143 AW768588 AA394195 AI288450 AW512564 AI452775 AI056520 AA468602 AA872566 AI434739
AA291838 AI948623 AW768614 AI374753 AIW068174 AA884908 AI199346 AI199347 W94946 AI159995 AA877642
AI280646 AI307610 AA403310 R08205 AW182123 AI000999 RZ7808 AW026571 D20816 AI560350 T27657 AW960271
AI174628 AI432042 AI424528 AA809562 T17342 AI783866
AI056548 AW409843 AW263540 AA723669 AA909334 AA156120 AA157141 AA156125 AW409866 W19499 AA157229
AW887435

TABLE 6:

5	Pkey:	Unique Eos probeset identifier number
	ExAccn:	Exemplar Accession number, Genbank accession number
	UnigeneID:	Unigene number
	Unigene Title:	Unigene gene title
	AUC1:	70 th percentile of average intensity (AI) for probeset at each of 2,6,15,24,48, and 96 hour timepoints minus 70 th percentile AI at 0 hrs, summed over 5 experiments.
10	AUC2:	AUC1/90 th percentile of AI for aorta, aortic valve, vein, and artery.

	Pkey	ExAccn	UnigeneID	UnigeneTitle	AUC1	AUC2
15	314941	AA515902	Hs.130650	ESTs	1038	9
	327414			predicted exon	303.2	30.3
	321911	AF026944	Hs.293797	ESTs	429.2	42.9
	331578	AI248482	Hs.249989	ESTs	677.4	10.3
	332466	AB018259	Hs.118140	KIAA0716 gene product	395.2	39.5
20	313513	AW298600	Hs.141840	ESTs, Weakly similar to S59501 Interfero	324	32.4
	320635	N50617	Hs.80506	small nuclear ribonucleoprotein polypept	394.8	39.5
	326230			predicted exon	357.2	35.7
	313556	AA628517	Hs.118502		433.6	12
	313665	AW751201	Hs.120932	ESTs	-83	0.5
25	324852	AI380792	Hs.135104	ESTs	348.2	34.8
	314372	AL040178	Hs.142003	ESTs, Weakly similar to The KIAA0149 gen	-49.2	0.5
	311877	AA084248	Hs.85339	G protein-coupled receptor 39	-1309	0.2
	322262	AA632012	Hs.188746	ESTs	-247.8	1
	312173	AI821409	Hs.304471	ESTs, Highly similar to AF116865 1 hedge	-1025.8	1
30	319795	AB037821	Hs.145858	protocadherin 10	203.6	5.2
	313350	AW591949	Hs.57958	ETL protein	183.8	18.4
	326759			predicted exon	1654.4	1.2
	300318	AW444502	Hs.256982	ESTs, Highly similar to AF116865 1 hedge	-346	1
	313978	AI870175	Hs.13957	ESTs	576.6	2.3
35	306840	AI077477	Hs.307912	EST	56.4	0.4
	310272	AF216389	Hs.148932	semaphorin Rs, short form	-127.6	0
	315044	BE547674	Hs.204169	ESTs	-102.6	0
	321325	AB033100	Hs.300646	KIAA protein (similar to mouse paladin)	1080.6	4.8
	303251	AF240635	Hs.115897	protocadherin 12	1270.8	5.3
40	302378	AL109712	Hs.296506	Homo sapiens mRNA full length insert cDN	915.8	15.8
	315060	AA551104	Hs.189048	ESTs, Moderately similar to ALUC_HUMAN I	1236.8	4.9
	332048	AW337575	Hs.201591	ESTs	522.6	4.7
	337214			predicted exon	269	26.9
	311598	AW023595	Hs.232048	ESTs	796.4	20.2
45	304782	AA582081		gbn32h08.s1 NCL_CGAP_Gas1 Homo sapiens	316.4	10.5
	312802	AA644669	Hs.193042	ESTs	349.6	7.6
	302680	AW192334	Hs.38218	ESTs	638.6	63.9
	317452	AA972965	Hs.135568	ESTs	360.8	36.1
	318558	AW402677	Hs.146381	RNA binding motif protein, X chromosome	700.2	6.6
50	312149	T90309	Hs.269651	ESTs	274.2	7.5
	319267	F11802	Hs.6818	ESTs	238.2	23.8
	321510	H75391	Hs.255748	ESTs	231.8	23.2
	326198			predicted exon	581.6	8.2
	315730	H25899	Hs.201591	ESTs	281.6	9.7
55	310442	AW072215	Hs.208470	ESTs	-213	0.3
	331237	W87874	Hs.25277	hypothetical protein FLJ21065	285	0.5
	300469	BE301708	Hs.233955	hypothetical protein FLJ20401	26.6	0.3
	338316			predicted exon	1494.2	34.7
	330968	R44557	Hs.23748	ESTs	975.8	1.8
60	331019	NM_006033	Hs.65370	lipase, endothelial	201.2	0.9
	331261	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (I		478.6 13.5
	301822	X17033	Hs.271986	Integrin, alpha 2 (CD49B, alpha 2 subuni	356.2	1.7
	325544			predicted exon	1014.6	9.4
	328700			predicted exon	627.4	62.7
65	322882	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	84.8	5.7
	336034			predicted exon	782.6	78.3
	316580	AA938198	Hs.146123	hypothetical protein FLJ12972	746.4	13.8
	309931	AW341683		gbtnd13d01.x1 Soares_NFL_T_GBC_S1 Homo s		134.8 13.5
	330692	R39288	Hs.6702	ESTs	137	13.7
70	319962	H06350	Hs.135058	Human DNA sequence from clone RP5-650E9		14.8 0.5
	338033			predicted exon	540.6	14
	314943	Y00272	Hs.184572	cell division cycle 2, G1 to S and G2 to	-494.8	1
	332640	BE568452	Hs.5101	protein regulator of cytokinesis 1	-600	1
	338158			predicted exon	311.2	31.1
75	327038			predicted exon	351.8	35.2

	302655	AJ227892	Hs.146274	ESTs	180.2	18
	327568			predicted exon	229	22.9
	324801	AW770553	Hs.14553	sterol O-acyltransferase (acyl-Coenzyme	161.2	16.1
	317850	AI681545	Hs.152982	hypothetical protein FLJ13117	-690	1
5	322818	AW043782	Hs.293618	ESTs	128.4	4.5
	324628	AI685464	Hs.292638	ESTs	170.2	17
	317224	X73608	Hs.93029	sparc/osteonectin, cwcr and kazal-like d	-80	0
	310955	AK76732	Hs.263912	ESTs	466.8	46.7
	315240	R36772	Hs.172618	KIAA1106 protein	277	27.7
10	338388			predicted exon	267.6	26.8
	338442			predicted exon	258	25.6
	318817	AW247252	Hs.75514	nucleoside phosphorylase	1247.8	24.2
	338645			predicted exon	206	20.6
	313135	N58907	Hs.162430	ESTs	204.8	20.5
15	324716	BE169746	Hs.12504	hypothetical protein DKFZp761D081	203.6	20.4
	330305			predicted exon	199.8	20
	308248	AI560919		gb:ta41g10.x1 NCL CGAP_Ut1 Homo sapiens		199.4 19.9
	308888	AI833240		gb:at76d10.x1 Barstead colon HPLR87 Homo	198.2	19.8
	315622	A1796144	Hs.258188	Homo sapiens cDNA FLJ11674 fs, clone HE	191.2	19.1
20	323675	R43240	Hs.272168	tumor differentially expressed 1	189.2	18.9
	312164	T91980	Hs.221074	ESTs	187.5	18.6
	300378	Z45270	Hs.235873	hypothetical protein FLJ22672	271.6	18.7
	317478	AI343569	Hs.107000	Homo sapiens mRNA for WDC146, complete c		187 18.7
	317559	AW452344	Hs.129977	ESTs	184.2	18.4
25	317207	AI873346	Hs.214505	ESTs	182.8	18.3
	334834			predicted exon	178.8	17.9
	320925	D62892		gb:HUM337C07B Clontech human aorta polyA		177.2 17.7
	303289	AL121460	Hs.272673	hypothetical protein FLJ20508	316.4	17.6
30	328548			predicted exon	174.6	17.5
	317108	AA884000	Hs.8173	hypothetical protein FLJ10803	172.4	17.2
	318013	AI188183	Hs.144078	ESTs	326	17.2
	314299	AW382682	Hs.154840	ESTs	170.8	17.1
	317702	AW173339	Hs.135665	ESTs	169.8	17
	316094	AW975920	Hs.283361	ESTs	169.4	16.9
35	323706	AA377578	Hs.65234	hypothetical protein FLJ20596	169.2	16.9
	325843			predicted exon	321.4	16.9
	316012	AA764950	Hs.119898	ESTs	1047.2	16.9
	309687	AW236154	Hs.77385	myosin, light polypeptide 6, alkali, smooth mu	168.2	16.8
	323329	AL134744	Hs.10852	ESTs	168	16.8
40	312853	W05086	Hs.114256	ESTs	167.4	16.7
	313070	AI422023	Hs.161338	ESTs	298.6	16.6
	314096	AW977642	Hs.291742	ESTs	165.6	16.6
	338728			predicted exon	165.4	16.5
	316609	AW292520	Hs.122082	ESTs	165	16.5
45	305999	AA888220		gb:oj15h01.s1 NCL CGAP_Kid5 Homo sapiens		164.6 16.5
	312642	AW052128		gb:wx26c02.x1 NCL CGAP_Kid11 Homo sapien		164 16.4
	339236			predicted exon	163.6	16.4
	317058	AI217713	Hs.147588	ESTs	161.8	16.2
	311137	AW207582	Hs.196042	ESTs	582.2	16.2
50	310178	AI938450	Hs.147482	ESTs	161.2	16.1
	320745	H51696	Hs.89278	hypothetical protein FLJ11186	161	16.1
	317336	AW014637	Hs.130212	ESTs	160	16
	309871	AW300366		gb:xs63b05.x1 NCL CGAP_Kid11 Homo sapien		159.8 16
	302038	AC004076	Hs.129709	Homo sapiens chromosome 19, cosmid R3021		159 15.9
55	332237	N52883	Hs.102676	EST	159	15.9
	312362	AW015994		gb:UH-B10p-abh-g-09-0-UI.s1 NCL CGAP_S	158.6	15.9
	331558	N62401	Hs.48531	EST	158.6	15.9
	316215	AI684535	Hs.200811	ESTs	158.4	15.8
	336059			predicted exon	157.4	15.7
60	302790	AJ245245		gb:Homo sapiens mRNA for immunoglobulin	155.8	15.6
	328418			predicted exon	153.8	15.4
	304228	AK000149	Hs.29493	hypothetical protein FLJ20142	153.6	15.4
	331608	AW273285	Hs.50802	ESTs	153	15.3
	338962			predicted exon	664.4	15.3
65	317959	AI204202	Hs.130264	ESTs	152.6	15.3
	336228			predicted exon	152.4	15.2
	313534	AW072916	Hs.78743	zinc finger protein 131 (clone pHZ-10)	152.2	15.2
	317404	AI806867	Hs.126594	ESTs	152.2	15.2
	311943	AI469911	Hs.26498	hypothetical protein FLJ21657	152	15.2
70	314680	AI247425	Hs.152182	ESTs	151.4	15.1
	331484	N29696	Hs.44076	EST	151.2	15.1
	338116			predicted exon	151.2	15.1
	329863			predicted exon	150.6	15.1
	315555	AW452886	Hs.239107	ESTs	149.6	15
75	317039	AA868583	Hs.126153	ESTs	149.6	15
	331138	R63816	Hs.28445	ESTs	149.6	15

	316561	AB17222	Hs.121655	ESTs	149.4	14.9
	328695			predicted exon	149.2	14.9
	302282	BE396283	Hs.173987	eukaryotic translation initiation factor	148.4	14.8
	318761	F11802	Hs.6818	ESTs	148.2	14.8
5	323709	AW297246	Hs.288546	Homo sapiens cDNA FLJ14190 (ts, clone NT	148	14.8
	310790	AW192063	Hs.248865	ESTs	147.8	14.8
	316833	AW292614	Hs.124367	ESTs	147.8	14.8
	323176	NM_007350	Hs.82101	pleckstrin homology-like domain, family	229	14.8
	324188	AW274439	Hs.252709	ESTs	147.6	14.8
10	317441	AA922798	Hs.196583	ESTs	147.4	14.7
	317584	AB25890	Hs.220513	ESTs	146.8	14.7
	321798	AB308206	Hs.181959	ESTs	146.8	14.7
	304363	AA208045		gb:zn77705.s1 Stratagene hNT neuron 937	146.6	14.7
	313952	F20956		gb:HSPO05390 HMG Homo sapiens cDNA clone	146.6 14.7	
15	301909	AI702609	Hs.15713	ESTs	263.8	14.7
	309196	AI904895	Hs.9614	nucleophosmin (nucleolar phosphoprotein	148.2	14.6
	321860	N47474	Hs.212631	ESTs	146.2	14.6
	330187			predicted exon	146	14.6
	323042	AA463571	Hs.172550	polypyrimidine tract binding protein (he	145.6	14.6
20	313636	AA262397	Hs.201366	ESTs	145.2	14.5
	302437	AB024729	Hs.227473	UDP-N-acetylglucosamine:alpha-1,3-D-mannosid	145	14.5
	318197	AA73096	Hs.133403	ESTs	144.8	14.5
	302749	M16951		gb:human lg mu-chain mRNA VDJA-region, 5	144.6	14.5
	322357	AI734258	Hs.245367	ESTs, Weakly similar to ALU1_HUMAN ALU S	144.6 14.5	
25	300391	AB27371	Hs.288839	hypothetical protein FLJ12178	144.4	14.4
	326077			predicted exon	144.4	14.4
	302004	Y18264	Hs.123094	sal (Drosophila)-like 1	144	14.4
	320668	AA805666	Hs.146217	Homo sapiens cDNA: FLJ23077 (ts, clone L	144	14.4
30	331212	T8693	Hs.226410	ESTs	144	14.4
	311268	AB69727	Hs.231859	ESTs	143.2	14.3
	305159	AA659166	Hs.275668	EST, Weakly similar to EF1D_HUMAN ELONGATIONF	143	14.3
	304510	AA457391	Hs.119122	ribosomal protein L13a	142.8	14.3
	320852	AA772920	Hs.303527	ESTs	142.8	14.3
35	330854	AW291944	Hs.122139	ESTs	142.8	14.3
	318275	AW449952	Hs.190125	basic-helix-loop-helix-PAS protein	142.6	14.3
	314992	AB248879	Hs.211286	ESTs, Weakly similar to 1207289A reverse	142.2	14.2
	322631	AA001697	Hs.293565	ESTs, Weakly similar to putative p150 [H	142.2	14.2
	332283	R40855	Hs.100839	EST	142	14.2
40	302894	AA719572	Hs.274441	Homo sapiens mRNA; cDNA DKFZp434N011 (tr	141.2 14.1	
	301808	R35391	Hs.252631	reticulum 3	141	14.1
	318608	AI204491	Hs.151502	ESTs	141	14.1
	316499	AW292947	Hs.122872	ESTs	140.8	14.1
	317011	AI248760	Hs.150276	ESTs	140.8	14.1
45	321840	N45600	Hs.46534	Homo sapiens mRNA; cDNA DKFZp434P0714 (f	140.8 14.1	
	327385			predicted exon	140.8	14.1
	331264	AA278898	Hs.225979	hypothetical protein similar to small G	140.8	14.1
	324545	AW501944	Hs.127243	Homo sapiens mRNA for KIAA1724 protein,	140.4	14
	312986	AA211586		gb:zn56d05.s1 Stratagene muscle 937209 H	140.2	14
50	316053	AA825814	Hs.149065	ESTs	140.2	14
	330723	BE247449	Hs.31082	hypothetical protein FLJ10525	140.2	14
	304876	AA595765		gb:nl28g06.s1 NCL_CGAP_AA1 Homo sapiens	139.8 14	
	311379	AW134766	Hs.202450	ESTs	139.8	14
	318265	AW019873	Hs.146840	ESTs	139.8	14
	324137	AA393127	Hs.222762	ESTs	139.8	14
55	328262			predicted exon	139.6	14
	322349	AK001279	Hs.180171	Homo sapiens cDNA FLJ10417 (ts, clone NT	139.4	13.9
	323504	AA280223	Hs.130865	ESTs	139.4	13.9
	304261	AA059387		gb:zf66d01.s1 Soares retina N2b4HR Homo	139.2	13.9
60	310489	AW451493	Hs.235516	hypothetical protein PRO2955	139.2	13.9
	335946			predicted exon	139.2	13.9
	318155	AI041546	Hs.132133	ESTs	138.8	13.9
	313796	AI797169	Hs.208486	ESTs	138.6	13.9
	333977			predicted exon	138.6	13.9
65	324845	AW969635	Hs.283718	ESTs	138.2	13.8
	331139	R65706		gb:yl16g12.s1 Soares placenta Nb2HP Homo	138.2	13.8
	331131	R54797		gb:yg87b07.s1 Soares Infant brain 1N1B H	669.6	13.8
	321250	H58539	Hs.151692	ESTs	138	13.8
	312498	AA668782	Hs.191284	ESTs, Weakly similar to ALU1_HUMAN ALU S	137.8 13.8	
	331252	W52470	Hs.34578	alpha2,3-sialyltransferase	137.8	13.8
70	337407			predicted exon	137.8	13.8
	303973	AW512014		gb:xc68a03.x1 NCL_CGAP_Lym12 Homo sapien	137.4 13.7	
	314582	AA412258	Hs.188817	ESTs	137.4	13.7
	327373			predicted exon	137.2	13.7
	323367	AA234591	Hs.304123	ESTs	136.6	13.7
75	316207	AA832065	Hs.120260	ESTs	136.4	13.6
	315231	AA705809	Hs.119922	ESTs	136.2	13.6

	318592	T39310	Hs.1139	cold shock domain protein A	136.2	13.6
	320906	AW969706	Hs.293332	ESTs	136.2	13.6
	328937			predicted exon	136.2	13.6
	329073			predicted exon	136.2	13.6
5	318231	AV659082	Hs.134228	ESTs	136	13.6
	311992	AL360200	Hs.114145	ESTs	135.8	13.6
	316497	AA766457	Hs.136849	ESTs	135.8	13.6
	317677	AA968594	Hs.127868	ESTs	135.8	13.6
	321680	W02848	Hs.93704	ESTs	135.8	13.6
10	326080			predicted exon	135.8	13.6
	330938	AF036943	Hs.172619	KIAA1106 protein	135.8	13.6
	306573	AL134878	Hs.119500	ribosomal protein, large P2	135.6	13.6
	307383	AI223207	Hs.147888	EST	135.6	13.6
	311114	AW449382	Hs.195297	ESTs	135.6	13.6
15	320579	R15138	Hs.165570	Homo sapiens clone 25052 mRNA sequence	135	13.5
	301328	AA884104	Hs.125546	ESTs	134.8	13.5
	312063	N58198	Hs.182898	ESTs	134.8	13.5
	323036	H09604	Hs.13268	ESTs	134.6	13.5
	332776	AF241850	Hs.151428	ret finger protein 2	134.4	13.4
20	332494	AA282330	Hs.145668	ESTs	134.2	13.4
	334376			predicted exon	134.2	13.4
	313264	N93416	Hs.118228	ESTs	133.6	13.4
	313669	AA351109	Hs.5437	Tax1 (human T-cell leukemia virus type I	133.2	13.3
	312083	T87398	Hs.205816	ESTs	132.6	13.3
25	319354	AA993807	Hs.167367	ESTs	132.6	13.3
	307414	AI242106		gbxh92a02.x1 Soares_NFL_T_GBC_S1 Homo s	132.2	13.2
	312771	AA018515	Hs.264482	App12 (autophagy 12, S. cerevisiae)-like	131.8	13.2
	313004	AI274963	Hs.145800	ESTs	131.2	13.1
	300995	AW510641	Hs.258018	ESTs	220.6	13
30	319323	F12650	Hs.13287	ESTs	125.4	12.5
	329451			predicted exon	123.4	12.3
	337603			predicted exon	572	12.2
	312480	R68651	Hs.144997	ESTs	121.4	12.1
	324934	AW452051	Hs.147546	ESTs	119.4	11.9
35	320723	BE178025	Hs.7942	hypothetical protein FLJ20080	117	11.7
	318188	AA792566		gbqj7402.y5 NCI_CGAP_Ov26 Homo sapiens	116.6	11.7
	320873	AF238869	Hs.283955	Homo sapiens clone GLSH-2 similar to gli	112.8	11.3
	331005	BE003191	Hs.119555	ESTs	112.8	11.3
	304969	AA614406		gbznp4805.s1 NCI_CGAP_Br11 Homo sapiens	112.4	11.2
40	319799	AI139253	Hs.227767	zinc finger protein 41	111.2	11.1
	302610	AA347945	Hs.256024	ESTs	111	11.1
	309485	AW130320	Hs.108124	ribosomal protein S4, X-linked	111	11.1
	311880	AW419225	Hs.256247	ESTs	110.2	11
	313981	AW452334	Hs.128148	ESTs	110.2	11
45	322442	W49701	Hs.29657	ESTs	109.4	10.9
	315099	AA806536	Hs.291841	ESTs	109	10.9
	304793	AA583264	Hs.182979	ribosomal protein L12	108.8	10.9
	330815	AA019211	Hs.238463	KIAA1238 protein	108.8	10.9
	304044	T81656	Hs.252259	ribosomal protein S3	714.8	10.8
50	325222			predicted exon	135	10.8
	325889			predicted exon	814.6	10.8
	321447	AW891130	Hs.38173	ESTs	107.8	10.8
	302990	AA496212	Hs.180182	ESTs	106.2	10.6
55	308106	AI476803		gbj77e12.x1 Soares_NSF_F8_9W_OT_PA_P_S	270.6	10.6
	310536	AI301041	Hs.150174	ESTs	106	10.6
	315257	AW157431	Hs.248941	ESTs	233	10.6
	318787	Z42313	Hs.22657	ESTs	105.8	10.6
	312306	AI927226	Hs.175610	ESTs	105.2	10.5
	326788			predicted exon	104.4	10.4
60	312234	AA830840	Hs.206934	ESTs	104	10.4
	314482	AW085525	Hs.134182	ESTs	234	10.4
	323597	AI185693	Hs.135119	ESTs	102.4	10.2
	302623	AW836724	Hs.194110	hypothetical protein PRO2730	162.4	10.2
	323594	AI791531	Hs.129993	ESTs	101	10.1
65	324315	N55761	Hs.194718	zinc finger protein 265	100.2	10
	314217	AA256465	Hs.188725	ESTs	99.2	9.9
	320932	AA554913	Hs.162287	ESTs	98.2	9.8
	327876			predicted exon	98.2	9.8
	319736	R17424	Hs.6650	vacuolar protein sorting 45B (yeast homo	98	9.8
70	327747			predicted exon	97.6	9.8
	327844			predicted exon	97.4	9.7
	318200	AI061192	Hs.166517	ESTs	97.2	9.7
	329414			predicted exon	97.2	9.7
	318296	AI089667	Hs.270713	ESTs	121.4	9.7
75	307010	AI140014		gbzqa6809.x1 Soares_fetal_heart_NbHH19W295	9.7	
	319792	AI138635	Hs.22968	ESTs	385.4	9.6

	305671	AA811688	Hs.62113	dUTPyrophosphatase	96	9.6
	329440			predicted exon	93.8	9.4
	310381	AI263059	Hs.145594	ESTs	93.4	9.3
	318824	F06771	Hs.27226	ESTs	93.4	9.3
5	328957			predicted exon	92.2	9.2
	318804	Z42549	Hs.160893	ESTs	92	9.2
	330836	AA055611	Hs.226568	ESTs, Moderately similar to ALU4_HUMAN A	92	9.2
	324592	AW752437	Hs.325708	ESTs	91.8	9.2
	311820	AW274545	Hs.254333	ESTs	91.4	9.1
10	321614	H86161		gb:ys94b01.r1 Soares retina N2bSHR Homo	91	9.1
	330306			predicted exon	91	9.1
	303096	AL080276	Hs.268562	regulator of G-protein signalling 17	90	9
	313275	AI027604	Hs.159650	ESTs	110.4	8.8
	302593	H54855	Hs.36958	ESTs	88	8.8
15	321421	BE465115	Hs.171688	ESTs	86.2	8.6
	330832	AI133530	Hs.62930	ESTs	456.4	8.6
	311847	AW301607	Hs.297260	ESTs	86	8.6
	322036	BE002723	Hs.301905	Homo sapiens cDNA FLJ14080 fls, clone HE	145.8	8.6
	328688			predicted exon	85.6	8.6
20	325251			predicted exon	85.4	8.5
	329088			predicted exon	85.4	8.5
	322524	W79027	Hs.271762	ESTs	84	8.4
	337953			predicted exon	451	8.3
25	323529	AA284397	Hs.201485	Homo sapiens clone FLC0664 PRO2866 mRNA	82.6	8.3
	307041	AI144243		gb:qb85b12.x1 Soares_fetal_heart_NbH119W	308.8	8.2
	318285	AI332454	Hs.158412	ESTs	81.4	8.1
	312021	AA759263	Hs.14041	ESTs	81	8.1
	329350			predicted exon	81	8.1
	326169			predicted exon	80.4	8
30	338038			predicted exon	1024.2	7.9
	312549	AI214510	Hs.146304	ESTs	77.4	7.7
	312542	D60076		gb:HUM084E10A Clontech human fetal brain	76.8	7.7
	320992	AB026891	Hs.225972	solute carrier family 7, (cationic amino	76	7.6
	318596	AI470235	Hs.172698	EST	150.8	7.5
35	315650	AA649042	Hs.269615	ESTs	73.4	7.3
	324328	AA447276	Hs.292020	ESTs	210.4	7.1
	332622	R10674	Hs.128856	CSR1 protein	70.2	7
	328229			predicted exon	69.4	6.9
40	319110	T75260	Hs.98321	hypothetical protein FLJ14103	68.6	6.9
	316133	AI187742	Hs.125562	ESTs	308.6	6.9
	303992	AW515800		gb:cd88g01.x1 NCLCGAP_GC6 Homo sapiens	67.8	6.8
	322675	AA017656	Hs.146580	enolase 2, (gamma, neuronal)	377.2	6.7
	325753			predicted exon	105.2	6.6
45	312539	AI004377	Hs.200360	Homo sapiens cDNA FLJ13027 fls, clone NT	92.2	6.4
	302592	AA294921	Hs.250811	v-rat simian leukemia viral oncogene hom	361.6	6.3
	314578	AA410183	Hs.137475	ESTs	201.6	6.1
	335986			predicted exon	108.6	6
50	321478	AW402593	Hs.123253	hypothetical protein FLJ22009	528	6
	305192	AA666019		gb:ag44a04.s1 Jia bone marrow stroma Hom	58.6	5.9
	304275	AA070605		gb:zm53h09.s1 Stratagene fibroblast (937	78.6	5.6
	302779	AI235667		gb:Homo sapiens mRNA for immunoglobulin	278.8	5.5
	301976	T97905	Hs.77256	enhancer of zeste (Drosophila) homolog 2	479.2	5.4
	316021	AW293399	Hs.144904	nuclear receptor co-repressor 1	792.4	5.3
55	320802	BE336699	Hs.185055	BENE protein	2423.8	5.3
	317282	AI733112	Hs.176101	ESTs	523.2	5.1
	316827	AI380429	Hs.172445	ESTs	578	5.1
	303190	BE280787	Hs.16079	hypothetical protein FLJ10233	223	5.1
	315587	AI268399	Hs.140489	ESTs	136.2	5
	333122			predicted exon	399	5
60	310214	AI220072	Hs.165893	ESTs	234.4	4.9
	320089	D43945	Hs.113274	transcription factor EC	68	4.9
	309328	AW024348	Hs.233191	EST, Weakly similar to A27217 glucose tr	258.8	4.8
	318971	Z44067	Hs.10957	ESTs	376.6	4.8
	327220			predicted exon	47.4	4.7
65	315757	AW014605	Hs.179872	ESTs	177.4	4.7
	320730	R68669	Hs.151072	ESTs	205.2	4.6
	313339	AI682536	Hs.163495	Homo sapiens cDNA FLJ13608 fls, clone PL	260	4.5
	318634	T49598	Hs.156832	ESTs	475.2	4.5
	320955	AW820035	Hs.278679	a disintegrin and metalloproteinase doma	388.6	4.4
70	306605	AI000497	Hs.119500	ribosomal protein largeP2	81.6	4.4
	308349	AW051913		gb:vw24a08.x1 NCLCGAP_Kid11 Homo sapion	102.4	4.3
	306004	AA889992	Hs.2186	eukaryotic translation elongation factor 1ga	451.2	4.2
	330020			predicted exon	61.2	4.1
	302308	AW327279	Hs.91379	ribosomal protein L26	342	3.9
75	314648	AW979268		gb:EST391378 MAGE resequences, MAGP Homo	56.4	3.8
	315131	AI753709	Hs.152484	ESTs	130.4	3.7

	313690	AA93591	Hs.78146	platelet/endothelial cell adhesion molec	3179.6	3.6
	333585			predicted exon	175.4	3.5
	312911	H93366	Hs.7587	Homo sapiens cDNA: FLJ21962 fls, clone H	219	3.5
	322966	AA633669	Hs.235920	Homo sapiens cell recognition molecule C	350.2	3.4
5	312492	R71072	Hs.191269	ESTs	322.8	3
	318988	Z44203	Hs.26418	ESTs	25	2.5
	332363	AI123705	Hs.106932	ESTs	773.4	2.5
	324181	AI025476	Hs.131628	ESTs	634.8	2.4
	311717	AW205389	Hs.312830	ESTs	54.2	2.4
10	321342	AA127984	Hs.222024	transcription factor BMAL2	23.4	2.3
	308852	AI829848	Hs.182937	peptidylprolyl isomerase A (cyclophilin A)	92	2.3
	331466	AA373210	Hs.43047	Homo sapiens cDNA FLJ13585 fls, clone PL	494	2.3
	320279	AB033062	Hs.134970	DKFZP434N178 protein	76.2	2.2
	322221	N24236	Hs.179662	nucleosome assembly protein 1-like 1	253.2	2.1
15	302925	AL137449	Hs.126666	homeo box B4	136.6	2.1
	331384	AB041035	Hs.93847	NADPH oxidase 4	720	1.8
	300938	AA514418	Hs.152320	ESTs, Weakly similar to 1605244A erythro	27	1.8
	312695	AW196863	Hs.200242	ESTs	303.8	1.6
	320223	W35132	Hs.267442	ESTs	189	1.5
20	332743	AW247977	Hs.87595	translocase of inner mitochondrial membr	14.4	1.4
	331039	AW376685	Hs.18625	Mitochondrial Acyl-CoA Thioesterase	529.8	1.4
	333123			predicted exon	396.2	1.4
	328455			predicted exon	91.8	1.3
	334458			predicted exon	406.4	1.3
25	313478	AA643008	Hs.192775	ESTs	413.4	1.1
	309899	AW338564	Hs.217493	annexin A2	-30.8	1
	311735	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fls, clone NT	-62.8	1
	312953	NM_001992	Hs.128087	coagulation factor II (thrombin) recepto	-73.6	1
	313055	AW367295	Hs.241175	ESTs	-43.8	1
30	313291	AI267970	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	-63	1
	315059	AW275110	Hs.271106	ESTs	-67	1
	322284	AI792140	Hs.49265	ESTs	-395.2	1
	322450	AL121278	Hs.25144	ESTs	-1.6	1
	324803	AW975183	Hs.292663	ESTs	4.4	1
35	331495	AW970939	Hs.291039	ESTs	-282.8	1
	333610			predicted exon	-152.6	1
	335093			predicted exon	-23.2	1
	339403			predicted exon	-331.2	1
40	302820	X04588	Hs.85844	neurotrophic tyrosine kinase, receptor,	591.2	1
	322770	R56151	Hs.93589	Homo sapiens mRNA; cDNA DKFZp564B1162 (f	276.6	1
	323755	AW300094	Hs.138252	ESTs	135	0.9
	326946			predicted exon	727.4	0.9
	315343	BE144306	Hs.179891	ESTs, Weakly similar to P4HA_HUMAN PROLY	122.8	0.9
45	311168	AK001270	Hs.196086	hypothetical protein FLJ10408	304	0.9
	329732			predicted exon	109.2	0.9
	321415	BE621807	Hs.3337	transmembrane 4 superfamily member 1	414.8	0.7
	333121			predicted exon	87.8	0.7
	333120			predicted exon	379.8	0.7
	330392	AW797956	Hs.75748	proteasome (prosome, macropain) subunit,	589.2	0.7
50	314711	AA769365	Hs.126058	ESTs	-87	0.6
	330865	BE409857	Hs.69499	hypothetical protein	347.4	0.6
	333169			predicted exon	-1182	0.6
	335095			predicted exon	106.4	0.6
	335815			predicted exon	-156	0.6
55	330232			predicted exon	102.6	0.6
	330823	AA031565	Hs.221255	ESTs, Moderately similar to ALU5_HUMAN A	-62	0.5
	331704	F04225	Hs.66032	ESTs	-14.8	0.5
	302842	NM_016428	Hs.130719	NESH protein	267.6	0.5
	304484	AA432067	Hs.258373	ESTs	85	0.5
60	310230	AK000377	Hs.144840	homolog of mouse C2PA	-70	0.4
	301531	AK077462	Hs.134084	ESTs	-195.4	0.4
	306337	AA954221	Hs.73742	ribosomal protein, large, P0	-33.4	0.4
	331327	N46436	Hs.109221	ESTs	-392	0.4
	332961			predicted exon	-5.6	0.4
65	322796	W31178	Hs.154140	Homo sapiens ovary-specific acidic prote	-880.6	0.3
	328857			predicted exon	55.2	0.3
	316342	AA743935	Hs.202329	ESTs	43.4	0.3
	331263	AW780192	Hs.267596	ESTs	-180.4	0.3
	335987			predicted exon	-134	0.3
70	311923	T60843	Hs.189679	ESTs	12.2	0.3
	310522	AW134529	Hs.244647	ESTs	-187.8	0.3
	315363	AA759190	Hs.121454	ESTs, Weakly similar to olfactory recept	80	0.3
	302032	NM_001992	Hs.128087	coagulation factor II (thrombin) recepto	-877	0.3
	313140	BE265133	Hs.217493	annexin A2	95.4	0.3
75	310860	AW015920	Hs.161359	ESTs	-239	0.3
	317899	AI952430	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	-715.2	0.3

	328520		predicted exon	-109.2	0.2
	302406	NM_012099	CD3-epsilon-associated protein; antisens	10	0.2
	311804	AI866921	Hs.203349 Homo sapiens cDNA FLJ12149 ts, clone MA	-252.6	0.2
	315065	AK001122	Hs.105859 hypothetical protein FLJ10260	-46.2	0.2
5	314129	AA228366	Hs.115122 ESTs	-308.8	0.2
	335697		predicted exon	-47.2	0.2
	335989		predicted exon	89	0.2
	320606	AW887943	Hs.127216 hypothetical protein FLJ13465	-205.6	0.2
	329745		predicted exon	103	0.2
10	313628	AW419069	Hs.209670 ESTs	-177.8	0.2
	334616		predicted exon	-836.6	0.2
	308820	AI821267	Hs.207243 EST	-7.2	0.2
	320416	AI026984	Hs.293662 ESTs	-18.4	0.2
	335211		predicted exon	-142	0.2
15	323629	AA375957	Hs.6682 ESTs	-100	0.1
	331420	AW452804	gb UJ-H-BI3-aly-h-11-0-ULs1 NCL_CGAP_Su	83	0.1
	315984	AI015862	Hs.131793 ESTs	-250.6	0.1
	332833		predicted exon	-374.2	0.1
	332607	NM_002314	Hs.36566 LIM domain kinase 1	-27.6	0.1
20	313467	AA004879	Hs.187820 ESTs	-288.2	0.1
	323333	AV651680	Hs.208558 ESTs	-735.6	0.1
	330775	AW247020	Hs.250747 SUMO-1 activating enzyme subunit 1	53.6	0.1
	333188		predicted exon	-1041.8	0.1
	332079	AI308876	Hs.103849 ESTs	19.4	0.1
25	322724	AF161442	Hs.191591 Homo sapiens HSPC324 mRNA, partial cds	-123.8	0.1
	303652	AI799111	Hs.64341 ESTs	-46.4	0.1
	303131	AW081061	Hs.103180 DC2 protein	-158.4	0.1
	320716	AI479439	Hs.171532 ESTs	-146.6	0.1
	300454	AA659037	Hs.163780 ESTs	-304	0.1
30	312757	AI285970	Hs.163817 ESTs	-445	0.1
	312391	RA3707	Hs.133159 ESTs, Weakly similar to PIHUSD salivary	-111.8	0.1
	308877	AI832519	gb aa69h03.x1 Barstead colon HPLRB7 Homo	-149.6	0
	311276	AI659166	Hs.207144 ESTs	-62.6	0
	302363	AW163799	Hs.198365 2,3-bisphosphoglycerate mutase	-15	0
35	321717	AW956580	Hs.42699 ESTs	-1059.8	0
	302638	AA463798	Hs.102696 MCT-1 protein	-332.2	0
	306352	AA961367	gb aa52a05.s1 NCL_CGAP_GC3 Homo sapiens	21.8	0
	313798	AI292148	Hs.71622 SWI/SNF related, matrix associated, acti	-87.2	0
	320807	AA135370	Hs.188536 Homo sapiens cDNA: FLJ21635 ts, clone C	-2222	0
40	320931	AW262836	Hs.252844 ESTs	-881.6	0
	332450	AW288085	Hs.11156 hypothetical protein	28.4	0
	332535	AF167706	Hs.19280 cysteine-rich motor neuron 1	-722	0
	335980		predicted exon	-421	0
	330746	AB033888	Hs.8619 SRY (sex determining region Y)-box 18	35.4	0
45	316820	AI627912	Hs.130783 Forssman synthetase	-373.6	0
	337429		predicted exon	-257	0
	331192	BE622021	Hs.152571 ESTs, Highly similar to IGF-II mRNA-bind	-33	0
	330609	AI346201	Hs.76118 ubiquitin carboxyl-terminal esterase L1	-280	0
	323593	AI739435	Hs.39168 ESTs	-3627.6	0
50	302704	AA531133	Hs.4253 hypothetical protein MGC2574	-278.6	0
	330534	NM_004579	Hs.82979 mitogen-activating protein kinase kinase	-244	0
	332374	X91195	Hs.100623 phospholipase C, beta 3, neighbor pseudo	-1204.2	0
	333221		predicted exon	-189.6	0
	335988		predicted exon	-122.6	0
55	330574	AI984144	Hs.66713 hepatitis delta antigen-interacting prot	-2257.4	0
	312052	BE821697	Hs.14317 nucleolar protein family A, member 3 (t-v	-359.2	0
	319568	AF131781	Hs.84753 hypothetical protein FLJ12442	-874.6	0
	337113		predicted exon	-24.6	0
	335149		predicted exon	-191.8	0

TABLE 6A

5 Table 6A shows the accession numbers for those pkeys lacking unigenelD's for Table 6. The pkeys in Table 7 lacking unigenelD's are represented within
 10 Tables 1-6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled
 using sequences derived from Genbank EST's and mRNAs. These sequences were clustered based on sequence similarity using Clustering and
 Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession"
 column.

15	Pkey:	Unique Eos probeset identifier number
	CAT number:	Gene cluster number
	Accession:	Genbank accession numbers
	Pkey	CAT Number Accession
20	320925	1525201_1 D62892 D79755 D62760
	321614	87866_1 H86161 AA054308 AA018955
	313952	136885_1 F20956 AA129374 AA133740 AW819878
	314648	293660_1 AW979268 AA878419 AA431342 AA431628
	302749	458_107 M16951 M16952 M16948 M16949 M16950
25	312362	764066_1 AWO15994 R39898 AWO00978 A1598202 A1521706
	312542	1522649_1 D60076 D60259 D61037
	312642	1005225_1 AW052128 H51439 H51481
	312986	171679_1 AA211586 F35799 AA211641 F29720 AW937387 AW937408
	329350	c_x_hs
30	329414	c_y_hs
	329440	c_y_hs
	329451	c_y_hs
	338033	CH22_6528FG_LINK_EM:AC00
	338038	CH22_6535FG_LINK_EM:AC00
35	338116	CH22_6650FG_LINK_EM:AC00
	338158	CH22_6700FG_LINK_EM:AC00
	329732	c14_p2
	329745	c14_p2
	308106	A1476803
40	329863	c14_p2
	338316	CH22_6944FG_LINK_EM:AC00
	308248	A1560919
	338388	CH22_7034FG_LINK_EM:AC00
	338442	CH22_7109FG_LINK_EM:AC00
	338645	CH22_7410FG_LINK_EM:AC00
45	338728	CH22_7527FG_LINK_EM:AC00
	308877	A1832519
	338962	CH22_7838FG_LINK_DJ32110
	308886	A1833240
50	333120	CH22_349FG_81_3_LINK_EM:A
	333121	CH22_350FG_81_4_LINK_EM:A
	333122	CH22_351FG_81_6_LINK_EM:A
	333123	CH22_352FG_81_7_LINK_EM:A
	333168	CH22_400FG_94_1_LINK_EM:A
	333169	CH22_401FG_94_2_LINK_EM:A
55	333221	CH22_458FG_105_1_LINK_EM:
	326077	c17_hs
	326080	c17_hs
	326169	c17_hs
	326198	c17_hs
60	326230	c17_hs
	333585	CH22_846FG_203_4_LINK_EM:
	333610	CH22_871FG_217_5_LINK_EM:
	335093	CH22_2423FG_492_3_LINK_EM
	335095	CH22_2425FG_492_5_LINK_EM
65	335149	CH22_2484FG_499_5_LINK_EM
	326759	c20_hs
	333977	CH22_1254FG_309_6_LINK_EM
	326788	c20_hs
	335211	CH22_2550FG_511_2_LINK_EM
70	305192	AA666019
	303973	AW512014
	303992	AW515800
	326946	c21_hs
75	328229	c_6_hs
	328262	c_6_hs

328418 c_7_hs
 328455 c_7_hs
 335697 CH22_3058FG_596_12_LINK_E
 328520 c_7_hs
 328548 c_7_hs
 335815 CH22_3187FG_618_3_LINK_EM
 328688 c_7_hs
 328695 c_7_hs
 307010 AI140014
 337113 CH22_5058FG_493_1_
 307041 AI144243
 328700 c_7_hs
 335946 CH22_3324FG_646_20_LINK_D
 335986 CH22_3366FG_654_10_LINK_D
 335987 CH22_3367FG_654_11_LINK_D
 335988 CH22_3368FG_654_12_LINK_D
 335989 CH22_3369FG_655_2_LINK_DJ
 335990 CH22_3370FG_655_4_LINK_DJ
 337214 CH22_5288FG_613_7_
 330020 c16_p2
 305989 AA888220
 328857 c_7_hs
 328937 c_8_hs
 328957 c_8_hs
 330187 c_4_p2
 337407 CH22_5607FG_755_1_
 337429 CH22_5633FG_762_3_
 330232 c_5_p2
 307414 AI242106
 330305 c_7_p2
 330306 c_7_p2
 337603 CH22_5896FG_LINK_C20H12
 337953 CH22_6395FG_LINK_EMAC00
 339236 CH22_8181FG_LINK_BA354I1
 339403 CH22_8384FG_LINK_BA232E1
 309349 AW051913
 325222 c10_hs
 325251 c10_hs
 318188 956161_1 AI792566 AI053836 AI054127 AI792489 AI288324
 309871 AW300366
 325544 c12_hs
 309931 AW341683
 332833 CH22_50FG_17_7_LINK_C20H1
 302779 33837_1 AJ235667 AJ235666 AJ235664 AJ235665 AJ235668 AJ235669 AJ235670
 302790 34166_1 AJ245245 AJ245247 AJ245257 AJ245254 AJ245256 AJ245253 AJ245203 AJ245250 AJ245252 AJ245243 AJ245204
 AJ245201 AJ245206 AJ245246 AJ245255 AJ245205 AJ245202 AJ245251 AJ245249 AJ245207 AJ245244
 332961 CH22_185FG_48_18_LINK_EM:
 325753 c14_hs
 327036 c21_hs
 325843 c16_hs
 325889 c16_hs
 304281 AA059387
 304275 AA070605
 334376 CH22_1670FG_379_8_LINK_EM
 327220 c_1_hs
 304363 AA206045
 334458 CH22_1757FG_391_2_LINK_EM
 327365 c_1_hs
 327373 c_2_hs
 334616 CH22_1923FG_411_15_LINK_E
 327414 c_2_hs
 327568 c_3_hs
 336034 CH22_3419FG_678_5_LINK_DJ
 336059 CH22_3445FG_684_2_LINK_DJ
 334834 CH22_2148FG_439_3_LINK_EM
 304782 AA582081
 304876 AA595755
 327747 c_5_hs
 336228 CH22_3626FG_730_4_LINK_DA
 329073 c_x_hs
 329088 c_x_hs
 304969 AA814406
 327844 c_5_hs
 327876 c_6_hs
 306352 AA961367
 331131 genbank_R54797 R54797

WO 02/079492

PCT/US02/04915

331139 genbank_R65706 R65706
331420 675963_1 AW452904 AW449414 BE467906 A1298565 BE549932 BE326357 F04362

TABLE 6B

5 Table 6B shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 6. The pkeys in Table 7 lacking unigeneID's are represented within Tables 1-6B. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

10 Pkey: Unique number corresponding to an Eos probe set
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA
 sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 15 Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
20	332961 Dunham, I. et al.	Plus	2521424-2521555
	333221 Dunham, I. et al.	Plus	3978070-3978187
	333585 Dunham, I. et al.	Plus	6234778-6234894
	333610 Dunham, I. et al.	Plus	6547007-6547116
	334376 Dunham, I. et al.	Plus	13902218-13902331
	334458 Dunham, I. et al.	Plus	14353496-14353572
25	334616 Dunham, I. et al.	Plus	15176123-15176470
	335149 Dunham, I. et al.	Plus	21497441-21497587
	335211 Dunham, I. et al.	Plus	21774611-21774680
	335697 Dunham, I. et al.	Plus	25481456-25481649
	335986 Dunham, I. et al.	Plus	27967791-27967852
30	335987 Dunham, I. et al.	Plus	27971413-27971481
	335988 Dunham, I. et al.	Plus	27977912-27978013
	335989 Dunham, I. et al.	Plus	27983788-27983860
	335990 Dunham, I. et al.	Plus	27988532-27988608
	336034 Dunham, I. et al.	Plus	29014404-29014590
35	337953 Dunham, I. et al.	Plus	6827029-6827125
	338033 Dunham, I. et al.	Plus	8092128-8092271
	338038 Dunham, I. et al.	Plus	8138219-8138392
	338316 Dunham, I. et al.	Plus	17089711-17089988
	338442 Dunham, I. et al.	Plus	19980640-19980698
40	338962 Dunham, I. et al.	Plus	29581892-29582020
	332833 Dunham, I. et al.	Minus	1119848-1119705
	333120 Dunham, I. et al.	Minus	3307508-3307427
	333121 Dunham, I. et al.	Minus	3308446-3308358
	333122 Dunham, I. et al.	Minus	3309596-3309531
45	333123 Dunham, I. et al.	Minus	3310817-3310749
	333168 Dunham, I. et al.	Minus	3729896-3729788
	333169 Dunham, I. et al.	Minus	3730864-3730767
	333977 Dunham, I. et al.	Minus	8722928-8722725
	334834 Dunham, I. et al.	Minus	17182681-17182535
50	335093 Dunham, I. et al.	Minus	21297367-21297214
	335095 Dunham, I. et al.	Minus	21292546-21292381
	335815 Dunham, I. et al.	Minus	26320518-26320421
	335946 Dunham, I. et al.	Minus	27487203-27487035
	336059 Dunham, I. et al.	Minus	29184079-29183969
55	336228 Dunham, I. et al.	Minus	30904602-30904497
	337113 Dunham, I. et al.	Minus	21233344-21233237
	337214 Dunham, I. et al.	Minus	26095902-26095502
	337407 Dunham, I. et al.	Minus	31886652-31886567
	337429 Dunham, I. et al.	Minus	32086238-32086079
60	337603 Dunham, I. et al.	Minus	1299296-1299194
	338116 Dunham, I. et al.	Minus	10614071-10613814
	338158 Dunham, I. et al.	Minus	11794465-11794343
	338388 Dunham, I. et al.	Minus	18662403-18662305
	338645 Dunham, I. et al.	Minus	24063839-24063775
65	338728 Dunham, I. et al.	Minus	25949039-25948927
	339236 Dunham, I. et al.	Minus	32773355-32773202
	339403 Dunham, I. et al.	Minus	34050728-34050625
	325222 6525287	Minus	22332-22473
	325251 6682448	Minus	411693-411751
70	325544 6682452	Plus	171228-171286
	325753 6682474	Plus	398512-398621
	329745 6065779	Plus	174774-175142
	329732 6065783	Plus	161252-161322
	329863 6691797	Plus	196801-196971
75	325889 5867087	Plus	223829-223891

WO 02/079492

PCT/US02/04915

	325843	6552453	Minus	7126-7232
	330020	6671887	Plus	172397-172491
	326198	5867215	Minus	80295-80674
	326230	5867230	Minus	301868-301972
5	326169	5867255	Minus	128321-128388
	328077	6682495	Minus	312108-312168
	328080	6682495	Plus	478644-478847
	326759	6249610	Plus	97216-97311
	326788	6682503	Plus	277132-277335
10	326946	6004446	Minus	116877-116967
	327036	6531965	Plus	319951-320040
	327220	5867525	Minus	65701-65781
	327365	6552412	Minus	118133-118198
	327414	5867750	Plus	102461-102586
15	327373	5867792	Minus	8188-8742
	327568	5867811	Minus	46152-46287
	330187	6706138	Plus	212923-213020
	327747	5867947	Plus	115322-115498
	327844	6249582	Minus	18895-18958
20	330232	6013528	Plus	113655-113830
	328229	5868105	Minus	120936-121053
	327876	5868140	Plus	103882-104034
	328262	6381906	Plus	11867-12027
	328688	5868262	Plus	626030-626094
25	328700	5868264	Plus	764089-764203
	328695	5868264	Plus	318632-318695
	328418	5868409	Minus	258811-258894
	328455	5868431	Plus	385576-385633
	328520	5868477	Plus	1942075-1942246
30	328548	5868487	Plus	72301-72397
	328857	6381927	Minus	80557-81051
	330305	4877982	Minus	52269-52365
	330306	4877982	Plus	96161-96233
	328937	5868500	Minus	1448241-1448333
35	328957	6456773	Plus	219195-219297
	329073	5868596	Plus	37838-37956
	329088	5868608	Plus	116738-116950
	329350	6456785	Plus	98911-98969
	329414	5868874	Plus	942555-942643
40	329440	5868885	Plus	21943-22063
	329451	5868887	Plus	25974-26048

TABLE 7:

5 Table 7 depicts Seq ID No., UnigeneID, UnigeneTitle, Pkey, and ExAccn for all of the sequences in Table 8. Seq ID No links the nucleic acid and protein sequence information in Table 8 to Table 7.

10	Pkey:	Unique Eos probaset identifier number		
	ExAccn:	Exemplar Accession number, Genbank accession number		
	UnigeneID:	Unigene number		
	Unigene Title:	Unigene gene title		
	Seq.ID.No.:	Sequence Identification Number found in Table 8		
15	PKey	ExAccn	Unigene ID	Unigene Title
				SEQ ID NO
	101545	BE246154	Hs.154210	endothelial differentiation, sphingolipid
	115819	AA486620	Hs.41135	endomucin-2
	424503	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,
	102917	AJ016712	Hs.287797	integrin, beta 1 (fibronectin receptor,
20	102915	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	105330	AW338625	Hs.22120	ESTs
	107385	NM_005397	Hs.16426	podocalyxin-like
	102024	AA301867	Hs.76224	EGF-containing fibulin-like extracellular
	102024	AA301867	Hs.76224	EGF-containing fibulin-like extracellular
25	134416	X68264	Hs.211579	melanoma cell adhesion molecule
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	104865	T79340	Hs.22575	B-cell CLL/lymphoma 6, member B (zinc fi
	106124	H93366	Hs.7567	Homo sapiens cDNA: FLJ21962 fis, clone H
	109001	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to
30	104764	AI039243	Hs.278585	ESTs
	133200	AB037715	Hs.183639	hypothetical protein FLJ10210
	105263	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	102892	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin
35	109456	AW956580	Hs.42699	ESTs
	110906	AA035211	Hs.17404	ESTs
	119073	BE245360	Hs.279477	ESTs
	132050	AI267615	Hs.38022	ESTs
	132490	NM_001290	Hs.4980	LIM domain binding 2
40	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	101714	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,
	133975	C18356	Hs.296944	tissue factor pathway inhibitor 2
	106793	H94997	Hs.16450	ESTs
	118511	N75620	Hs.43157	ESTs
45	101447	M21305		gbHuman alpha satellite and satellite 3
	314941	AA515902	Hs.130650	ESTs
	332466	AB018259	Hs.118140	KIAA0716 gene product
	313513	AW298600	Hs.141840	ESTs, Weakly similar to S59501 Interfero
	313556	AA628517	Hs.118502	ESTs
	313665	AW751201	Hs.51233	ESTs
50	314372	AL040178	Hs.142003	ESTs
	429276	AF056085	Hs.198612	G protein-coupled receptor 51
	101345	NM_005795	Hs.152175	calcitonin receptor-like
	418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecu
	103850	AA187101	Hs.213194	hypothetical protein MGC10895
55	133260	AA403045	Hs.6906	Homo sapiens cDNA: FLJ23197 fis, clone R
	101097	BE245301	Hs.89414	chemokine (C-X-C motif), receptor 4 (fus
	104786	AA027167	Hs.10031	KIAA0955 protein
	132173	X89426	Hs.41716	endothelial cell-specific molecule 1
	100420	D86983	Hs.118893	Melanoma associated gene
60	111018	AI287912	Hs.3628	mitogen-activated protein kinase kinase
	108507	AI554545	Hs.68301	ESTs
	104894	AF065214	Hs.18858	phospholipase A2, group IVC (cytosolic,
	118511	N75620	Hs.43157	ESTs
	125609	AA868063	Hs.104576	carbohydrate (keratan sulfate Gal-6) sul
65	101543	M31166	Hs.2050	perlecan-related gene, rapidly induced b
	102241	NM_007351	Hs.268107	multimerin
	101560	AW958272	Hs.347326	intercellular adhesion molecule 2
	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	105826	AA478756	Hs.194477	E3 ubiquitin ligase SMURF2
70	102804	NM_002318	Hs.83354	lysyl oxidase-like 2
	131647	AA359615	Hs.30089	ESTs
	103095	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and
	103037	BE018302	Hs.2894	placental growth factor, vascular endoth
	100405	AW291587	Hs.82733	nidogen 2
75	102012	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas

	101261	D30857	Hs.82353	protein C receptor, endothelial (EPCR)	Seq ID 120 & 121
	105729	H46612	Hs.293815	Homo sapiens HSPC265 mRNA, partial cds	Seq ID 122 & 123
	107216	D51069	Hs.211579	melanoma cell adhesion molecule	Seq ID 124 & 125
	131080	NM_001955	Hs.2271	endothelin 1	Seq ID 126 & 127
5	131486	F06972	Hs.27372	BMX non-receptor tyrosine kinase	Seq ID 128 & 129
	134299	AW580939	Hs.97199	complement component C1q receptor	Seq ID 130 & 131
	134983	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p	Seq ID 132 & 133
	115827	AA428000	Hs.283072	actin related protein 2/3 complex, subun	Seq ID 134 & 135
	133614	NM_003003	Hs.75232	SEC14 (S. cerevisiae)-like 1	Seq ID 136 & 137
10	116483	AJ348201	Hs.76118	ubiquitin carboxyl-terminal esterase L1	Seq ID 138 & 139
	132546	M24263	Hs.168383	Intercellular adhesion molecule 1 (CD54)	Seq ID 140 & 141
	133678	AW247252	NA	nucleoside phosphorylase	Seq ID 142 & 143
	130184	H58306	Hs.15165	retinoic acid induced 14	Seq ID 144 & 145
	134786	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous	Seq ID 146 & 147
15	129371	X06828	Hs.110802	von Willebrand factor	Seq ID 148 & 149
	418506	AA084248	Hs.85339	G protein-coupled receptor 39	Seq ID 150 & 151
	322262	AA632012	Hs.188748	ESTs	Seq ID 152 & 153
	312173	AJ821409	Hs.304471	EST	Seq ID 154 & 155
	319795	AB037821	Hs.146858	protocadherin 10	Seq ID 156 & 157
20	313978	AJ870175	Hs.13957	ESTs	Seq ID 158 & 159
	306840	AJ077477	Hs.307912	ESTs	Seq ID 160 & 161
	310272	AF216389	Hs.148932	sema domain, transmembrane domain (TM),	Seq ID 162 & 163
	310272	AF216389	Hs.148932	sema domain, transmembrane domain (TM),	Seq ID 164 & 165
	315044	BE547674	Hs.204169	ESTs, Weakly similar to S65657 alpha-1C-	Seq ID 166 & 167
25	321325	AB033100	Hs.300646	KIAA1274 protein (similar to mouse palad	Seq ID 168 & 169
	321325	AB033100	Hs.300646	KIAA1274 protein (similar to mouse palad	Seq ID 170 & 171
	303251	AF240635	Hs.115897	protocadherin 12	Seq ID 172 & 173
	302378	AL109712	Hs.296506	Homo sapiens mRNA full length insert cDN	Seq ID 174 & 175
	319267	F11802	Hs.6818	ESTs	Seq ID 176 & 177
30	310442	AW072215	Hs.208470	ESTs	Seq ID 178 & 179
	300469	BE301706	Hs.233955	hypothetical protein FLJ20401	Seq ID 180 & 181
	331237	W87874	Hs.25277	Homo sapiens cDNA FLJ10717 fls; clone NT	Seq ID 182 & 183
	330988	R44557	Hs.23748	ESTs	Seq ID 184 & 185
	301822	X17033	Hs.271986	Integrin, alpha 2 (CD49B, alpha 2 subuni	Seq ID 186 & 187
35	422573	AW297985	Hs.295726	Integrin, alpha V (vitronectin receptor	Seq ID 188 & 189
	133061	AJ188431	Hs.286638	prostate differentiation factor	Seq ID 190 & 191
	135235	AW298244	Hs.266195	ESTs	Seq ID 192 & 193
	101192	BE247295	Hs.78452	solute carrier family 20 (phosphate tran	Seq ID 194 & 195
	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	Seq ID 196 & 197
40	101741	NM_003199	Hs.326198	transcription factor 4	Seq ID 198 & 199
	321911	AF026944	Hs.293797	ESTs	Seq ID 200 & 201
	320535	N50617	Hs.80506	small nuclear ribonucleoprotein polypept	Seq ID 202 & 203
	326230			NM_017643:Homo sapiens hypothetical prot	Seq ID 204 & 205
45	132968	AF234532	Hs.61638	myosin X	Seq ID 206 & 207
	135073	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f	Seq ID 208 & 209
	108937	AL050107	Hs.24341	transcriptional co-activator with PDZ-bi	Seq ID 210 & 211
	116430	AK001531	Hs.66048	hypothetical protein FLJ10669	Seq ID 212 & 213
	104877	AJ138635	Hs.22858	Homo sapiens clone IMAGE:451939, mRNA se	Seq ID 214 & 215
	122697	AA420663	Hs.98321	hypothetical protein FLJ14103	Seq ID 216 & 217
50	112522	R68857	Hs.265499	ESTs	Seq ID 218 & 219
	304782	AA582081		gb:nn32h08.s1 NCL_CGAP_Gas1 Homo sapiens	Seq ID 220 & 221
	312802	AA644669	Hs.193042	ESTs	Seq ID 222 & 223
	302680	AW192334	Hs.38218	ESTs	Seq ID 224 & 225
	326198			Phase 2 & 3 Exons	Seq ID 226 & 227
55	331019	NM_006033	Hs.65370	lipase; endothelial	Seq ID 228 & 229

TABLE 8

Seq ID NO: 1 DNA sequence
 Nucleic Acid Accession #: NM_001400
 Coding sequence: 244-2208 (underlined sequences correspond to start and stop codons))

```

10 1      11      21      31      41      51
    |      |      |      |      |      |
    GTCGGGGGCA GCAGCAAGAT GCGAAGCGAG CGGTACAGAT CCGGGGCTCT CGAAAGCGAA 60
    CTTGCGCCCTG CTTGAGCGAG GCTGCGGTTT CGAGGGCCCT CTCAGGCCAA GGAAGGCTA 120
    CACAAAAGAGC CTGGATCACT CATGGAACCA CCGCTGAAGC CAGTGAAGGC TCTCTGCGCT 180
    CGCCCTCTAG CGTGTGCTG GAGTAGCGCC ACCCGGCTT CCGGGGACA CAGGGTTGGC 240
    ACCATGGGGC CACCAAGGT CCGGCTGGTC AAGGCCACCC GCAGCTGGT CTCTGACTAC 300
    GTCAACTATG ATATCATGCT CCGGCTTAC AACTACACGG GAAAGCTGAA TATCAGCGCG 360
    GACAAGGAGA ACAGCATTAA ACTGACCTCG GTGGTGTTC TTCTCATCTG CTGCTTTATC 420
    ATCTCTGAGA ACATCTTTGT CTTGCTGACC ATTTGGAATA CCAAGAAAT CCACCGACCC 480
    ATGTACTATT TTATTGGCAA TCTGGCCCTC TCAGACCTGT TGGCAGGAGT AGCCTACACA 540
    GCTAACCTGC TCTTGTCTGG GGGCACCACC TACAAGCTCA CTCGCGCCCA GTGGTTTCTG 600
    CGGGAAGGGA GTATGTTTGT GGCCTGTGCA GCTTCGCTGT TCAGTCTCCT CGCCATCGCC 660
    ATTGAGCGCT ATATCACAAT GCTGAAAATG AACTTCCACA ACGGGAGCAA TAACCTCCGC 720
    CTCTTCCGCG TAATCAGCGC CTGCTGGGTC ATCTCCCTCA TCCTGGGTGG CCGCTATC 780
    ATGGGCTGGA ACTGCATCAG TGGGCTGTCC AGCTGCTCCA CGGTGCTGCC GCTCTACCAC 840
    AAGCACTATA TCTCTTCTG CACCAAGGTC TTCACTCTGC TTCTGCTCTC CATCGTCATT 900
    CTGTACTGCA GAATCTACTC CTTGCTCAGG ACTCGGAGCC GCGGCTGAC GTTCCGCAAG 960
    AACATTCCA AGGCCAGCGC CAGCTCTGAG AAGTGGCTGG CGCTGCTCAA GACCGTAATT 1020
    ATCTCTCTGA GCGTCTTCAT CGCTGCTGG GCACCGCTCT TCATCTGCT CTGCTGGAT 1080
    GTGGGCTGCA AGGTGAAGAC CTGTGACATC CTCTTCAGAG CGGAGTACTT CCGGTGTTA 1140
    GCTGTGCTCA ACTCGGCAC CAACCCCATC ATTTACACTC TGACCAACAA GGAGATGCGT 1200
    CGGCCCTTCA TCGGATCAT GTCTGTCTGC AAGTGCCGGA GCGGAGACTC TGCTGGCAAA 1260
    TTCAGCGCAC CCATCATCGC CGGCTGGAA TTCAGCGCA GCAATCGGA CAATCTCTCC 1320
    CACCCCCAGA AAGACGAAGG GGACAACCCA GAGACCATTA TGCTTCTGG AAACGTCAC 1380
    TCTTCTCTCT AGAATCGGAA GCTGTCCACC CACCGGAAGC GCTCTTACT TGGTCGCTGG 1440
    CCACCCCTGA GTTGGGAJAA AATCTCTGG GCTTCGACTG CTGCCAGGGA GGAGCTGCTG 1500
    CAAGCCAGAG GGAGGAAGGG GGAGAAATAG AACAGCTGG TGGTGTGGGG TGTGTGGGG 1560
    TAGAGTATGT TCTGTGAAC AATGCACTGG GAAGGTGGA GATCAGTCC CGGCTGGAA 1620
    TATATATTTT ACCCCCTGG AGCTTTGATT TTGCACTGAG CCAAGGCTCT AGCATTTGCA 1680
    AGCTCTCTAAA GGGTTCATTT GGCCTCTCT CAAAGACTAA TGTCCTCATG TGAAGGCTC 1740
    TCTTGTCTG GAGCTTTGAG GAGATGTTT CTTTCACTTT AGTTTCAAAC CCAAGTGAGT 1800
    GTGTGCACTT CTGCTTCTTT AGGATGCCCC TGTACATCCC ACACCCCAAC CTCCTTCCC 1860
    TTCAATACCC TCCTCAACGT TCTTTTACTT TATACITTA CTACCTGAGA GTTATCAGAG 1920
    CTGGGGTGT GGAATGATCG ATCATCTATA GCAATAGGC TATGTTGAGT ACGTAGGCTG 1980
    TGGAAGATG AAGATGTTT GGAGGTGTAA AACAAATGCC TTGCTGAGG CCAAGTTTC 2040
    CATGTAAAGG GGATCCGTTT TTGGGAATTT GGTGGAAGTC ACTTTGATTT CTTTAAAAAA 2100
    CATCTTTTCA ATGAAATGTG TTACCAATTC ATATCCATTG AAGCCGAAAT CTGCATAAGG 2160
    AAGCCCACTT TATCTAAATG ATATTAGCCA GGATCCTTGG TGTCCTAGGA GAAACAGACA 2220
    AGCAAAACAA AGTGAACACC GAATGGATTA ACTTTTGCAA ACCAAGGGAG ATTTCTTAGC 2280
    AATGAGTCTT AACAAATATG ACATCCGTCT TTCCCACTTT TGTGTATGTT TATTTCAGAA 2340
    TCTTGTGTGA TTCAATTTCA GCAACAACAT GTTGTATTTT GTTGTGTAA AAGTACTTTT 2400
    CTTGATTTTT GAATGTATTT GTTTCAGGAA GAAGTCATTT TATGGATTTT TCTAACCCGT 2460
    GTTAACTTTT CTAGAATCCA CCTCTCTGTG CCTTAAAGCA TTACTTTAAC TGGTAGGGAA 2520
    CGCCAGAACT TTTAAGTCCA GCTATTCATT AGATAGTAAT TGAAGATATG TATAAATATT 2580
    ACAAGAATA AAAATATATT ACTGTCTCTT TAGTATGGTT TTCAGTCAA TTAACCGAG 2640
    AGATGTCTTG TTTTITTTAA AAGAAATAGT TTAATAGGT TTCTGACTTT TGTGGATCAT 2700
    TTTGCACATA GCTTTATCAA CTTTAAACA TTAATAAAT GATTTTTTTA AAG
  
```

Seq ID NO: 2 Protein sequence
 Protein Accession #: NP_001391

```

65 1      11      21      31      41      51
    |      |      |      |      |      |
    MGPTSVPLVK AHRSSVSDYV NYDIIVRHYN YTGKLNISAD KENSIKLTST VPILICCFII 60
    LENIFVLLTI WTKTKPHRPM YYPIGNLALS DLLAGVAYTA NLLSGATTY KLTFAQWFLR 120
    EGSMPVALSA SVFSLLAIAI ERYITMLKMK LHNGSNNFRL FLLISACNVI SLILGGLPIM 180
    GWNCSIALSS CSTVLPYHK HYILFCTTVP TLLLSIVIL YCRIYSLVRT RSRRLTFRKN 240
    ISKARSSEK SLALLKTVII VLSVFIACWA PLFILLLLDV GCKVKTCDIL FRAEYFLVLA 300
    VLSGNTNPII YTLTNKEMRR AFIRIMSCCK CPSSGASGKF KRPIIAGMEF SRSKSDNSSH 360
    PQKDEGDNPE TMSSGNVNS SS
  
```


5

Seq ID NO: 3 Nucleotide sequence:
 Nucleic Acid Accession #: NM_016242
 Coding sequence: 79-864 (underlined sequences correspond to start and stop codons))

```

15 1      11      21      31      41      51
    |      |      |      |      |      |
    AAGGCCCTGC CAGCTTGGGA GGGAAATTGTC CCTGCCTGCT TCTGGAGAAA GAAGATATTG 60
    ACACCATCTA CGGGCACCAT GGAAGTGTCT CAAGTGACCA TTCTTTTCT TCTGCCAGT 120
    ATTTCAGCA GTAACAGCAT AGGTGTTTGA GAGGCAGCTA ATAATTCCT TGTGTGTTACT 180
    ACAACAAAC CATCTATAAC AACACAAAC ACAGAATCAT TACGAAAAA TGTGTGTCACA 240
    CCAACACTGT GAACACTCC TAAAGGAACA ATCACCATTG AATTACTTAA AATGTCCTCTG 300
    ATGTCAACAG CTACTTTTT AACAACTAAA GATGAAGGAT TGAAAGCCAC AACCCTGAT 360
    GTCAGGAAGA ATGACTCCAT CATTTCAAA GTAACAGTAA CAAGTGTTAC ACTTCCCAAT 420
    GCTGTTTCAA CATTACAAAG TTCCAAACCC AAGACTGAAA CTCAGAGTTC AATTAAAAACA 480
    ACAGAAATAC CAGGTAGTGT TCTACAACCA GATGCATCAC CTCTAAAAC TGGTACATTA 540
    25 ACCTCAATAC CAGTTACAT TCCAGAAAAC ACCTCACAGT CTCAGTAAT AGACACTGAG 600
    GGTGGAAGAA ATGCAAGCAC TTCAGCAACC AGCCGGTCTT ATTCCAGTAT TATTTGCGG 660
    GTGGTTATTG CTTTGATTGT AATAACACT TCAGTATTGT TTCTGGTGGG TTTGTACCGA 720
    ATGTGCTCGA AGGCAGATCC GGGCACACCA GAAATGGAA ATGATCAACC TCAGTCTGAT 780
    30 AAAGAGAGCG TGAAGCTTCT TACCGTTAAG ACAATTCTCT ATGAGTCTGG TGAGCACTCT 840
    GCACAGAGAA AAACCAAGAA CTGACAGCTT GAGGAATTCT CTCCACACCT AGGCAATAAT 900
    TAGCTTTAAT CTTACAGTTC TATGCACCAA GCGTGGAAAA GGAGAAAGTC CTGCAGAATC 960
    AATCCCGACT TCCATACCTG CTGCTGG
  
```

35

Seq ID NO: 4 Protein sequence:
 Protein Accession #: NP_057326

```

40 1      11      21      31      41      51
    |      |      |      |      |      |
    MELIQTILF LLPSICSSNS TGVLEAANNS LVTITTKPSI TTPNTESLQK NVVPTPTGTT 60
    PKGTINELL KMSLMSTATF LTKDEGLKA TTDVVRKND IISNVTVTSV TLPNAVSTLQ 120
    SSKPKTETQS SIKTTEIPGS VLQPDASPSK TGTLTSPVT IPENTSGSQV IDTEGGKNAS 180
    TSATSRYSYS IILPVVIALI VITLSVFFLV GLYRMCWKAD PGTPEGNNDQ PQSDKESVKL 240
    45 LTVKTSIHES GEHSAQKTK N
  
```

Seq ID NO: 5 Nucleotide sequence:
 Nucleic Acid Accession #: NM_002205
 Coding sequence: 24..3173 (underlined sequences correspond to start and stop codons))

```

55 1      11      21      31      41      51
    |      |      |      |      |      |
    CAGGACAGGG AAGAGCGGGC GCTATGGGGA GCGGAGCGCC AGAGTCCCTT CTCCAGCGCG 60
    TGCAGCTGCG CTGGGGCCCC CGGCGCGGAC CCCGCTGCT GCGGCTGCTG TTGCTGCTCG 120
    TGCCGCGGCC ACCCAGGGTC GGGGGCTTCA ACTTAGACGC GGAGGCCCCA GCAGTACTCT 180
    CGGGGCCCCC GGGCTCCTTC TTGGGATTCT CAGTGGAGTT TTACGGCGCG GGAACAGACG 240
    GGTGCTAGGT GCTGGTGGGA GCACCCAAGG CTAATACCA GACGCCAGGA GTGCTGCAGG 300
    60 GTGGTCTGCT CTACCTCTGT CCTTGGGGTG CCAGCCCCAC ACAGTGACCC CCATTGAAT 360
    TTGACAGCAA AGGCTCTCGG CTCTGGAGT CCTCACTGTC CAGCTCAGAG GGAGAGGAGC 420
    CTGTGGAGTA CAAGTCTCTG CAGTGGTTGG GGGCAACAGT TCGAGCCCAT GGCTCCTCCA 480
    TCTTGGCATG CGCTCCACTG TACAGCTGGC GCACAGAGAA GGAGCCACTG AGCGACCCCG 540
    TGGGCACCTG CTACCTCTCC ACAGATAACT TCACCGAAT TCTGGAGTAT GCACCTTGCC 600
    65 GCTCAGATTG CAGCTGGGCA GCAGGACAGG GTTACTGCCA AGGAGGCTTC AGTCCCGAGT 660
    TCACCAAGAC TGGCCGTGTG GTTTTAGGTG GACCAGGAAG CTAATTCCTG CAAGGCCAGA 720
    TCCTGTCTGC CACTCAGGAG CAGATTGCAG AATCTTATTA CCCCAGTAC CTGATCAACC 780
    TGGTTCAGGG GCAGCTGCAG ACTCGCCAGG CCAGTTCCAT CTATGATGAC AGCTACCTAG 840
    GATACTCTGT GGCTGTTGGT GAATTCAGTG GTGATGACAC AGAAGACTTT GTTGTGGTGT 900
    70 GCCCCAAGG GAACCTCACT TACGGCTATG TCAACATCTT TAATGGCTCA GACATTGAT 960
    CCTCTACAA CTTCTCAGGG GAACAGATGG CCTCTACTT TGGCTATGCA GTGGCGGCGA 1020
    CAGACGTCAA TGGGAGCGGG CTGATGACT TGCTGTGGGG GGCACCCCTG CTCATGGATC 1080
    GGACCCCTGA CGGCGCGCT CAGGAGGTGG GCAGGGTCTA CGTCTACCTG CAGCACCCAG 1140
    75 CCGCATAGA GCCCACGCCC ACCCTTACCC TCACTGGCCA TGATGAGTTT GGCCGATTTG 1200
    CGAGCTCCTT GACCCCTCTG GGGGACCTGG ACCAGGATGG CTACAATGAT GTGGCCATCG 1260
    GGGCTCCCTT TGGTGGGGAG ACCCAGCAGG GAGTAGTGTT TGTATTTCCT GGGGGCCAG 1320
  
```

5 GAGGGCTGGG CTCTAAGCCT TCCCAGGTTT TGCAGCCCCCT GTGGGSCAGCC AGCCACACCC 1380
 CAGACTTCTT TGGCTCTGCC CTTCGAGGAG GCCGAGACCT GGATGGCAAT GGATATCTCT 1440
 ATCTGATTGT GGGGTCTCTT GTGTGTGGACA AGGCTGTGGT ATACAGGGGC CGCCCCATCG 1500
 TGTTCGCTAG TGGCTCCCTC ACCATCTTCC CGGCCATGTT CAACCCAGAG GAGCGGAGCT 1560
 GCAGCTTAGA GGGGAACCTT GTGGCTTGCA TCACCTTAGT CTCTGCTCTC AATGCTTCTG 1620
 GAAAACACGT TGTGACTCC ATTTGGTTTCA CAGTGGAACT TCAGCTGGAC TGGCAGAAGC 1680
 AGAAGGGAGG GGTACGGCGG GCACTGTCTC TGGCTCCAG GCAGGCAACC CTGACCCAGA 1740
 CCTGTCTCAT CCAGAATGGG GCTCGAGAGG ATTGCAGAGA GATGAAGATC TACCTCAGGA 1800
 10 ACAGTTCAGA ATTTGAGAGC AAATCTCTGC CGATTTCAT CGCTCTCAAC TTCTCTCTGG 1860
 ACCCCCAAGC CCCAGTGGAC AGCCACGGCC TCAGGCCAGC CCTACATTAT CAGAGCAGA 1920
 GCGGATAGA GGACAGGCT CAGATCTTGC TGGACTGTGG AGAAGACAAC ATCTGTGTGC 1980
 CTGACCTGCA GCTGGAAGTG TTTGGGGAGC AGAACCATGT GTACCTGGGT GACAAGAATG 2040
 CCTTGAAGCT CACTTTCCAT GCCCAGAATG TGGGTGAGGG TGGCGCTAT GAGGCTGAGC 2100
 TTTGGGTGAC CGCCCTTCCA GAGGCTGAGT ACTCAGGACT CGTCAGACAC CCAGGGAAGT 2160
 15 TCTCCAGGCT GAGCTGTGAC TACTTTGCCG TGAACAGAG CGCCTGCTG GTGTGTGACC 2220
 TGGGCAACCC CATGAAGGCA GGAGCCAGTC TGTGGGGTGG CCTTCGGTTT ACAGTCCCTC 2280
 ATCTCCGGGA CACTAAGAAA ACCATCCAGT TTGACTTCCA GATCCTCAGC AAGAATCTCA 2340
 ACATCTGCA AAGCAGCTG GTTCTCTTTC GGCTCTCGGT GGAGGCTCAG GCCCAGGTCA 2400
 CCTTGAACGG TGTCTCCAAG CTTGAGGCGG TGCTATTCCC AGTAAAGGAC TGGCATCCCT 2460
 20 GAGACCAACC TCAGAAGGAG GAGGACCTGG GACCTGTGCT CCACCATATC TATGAGCTCA 2520
 TCACCAAGG CCCAGCTCC ATTAGCCAGG GTGTGCTGGA ACTCAGCTGT CCCAGGCTC 2580
 TGAAGGTC AAGCTCCCTA TATGTAGCCA GAGTTAAGGG ACTCACTGC ACCACCAATC 2640
 ACCCATATTAA CCCAAGGGC CTGGAGTTGG ATCCCGAGGG TTCCCTGCAC CACAGCAAA 2700
 25 AACGGGAAGC TCCAAGCGC AGCTCTGCTT CCTCGGGACC TCAGATCCTG AAATGCCCGG 2760
 AGCTGAGTGT TTTGAGGCTG CGCTGTGAGC TCGGGGCCCT GCACCAACAA GAGAGCCAAA 2820
 GTCTGCAGTT GCATTTCGGA GTCTGGGCCA AGACTTTCTT GCAGCGGGAG CACAGCCAT 2880
 TTAGCTGCA GTGTGAGGCT GTGTACAAAG CCCTCAAGAT GCCCTACCGA ATCCTGCTC 2940
 GGCAGCTGCC CCAAAAGAG CGTCAGGTGG CCACAGCTGT GCATGGGACC AAGCAGAG 3000
 30 GCAGCTATGG CGTCCCACTG TGGATCATCA TCCTAGCCAT CTTGTTTGGC CTCTGCTCC 3060
 TAGGTCTACT CATCTACATC CTCTACAGC TTGATTTCTT CAAACGCTCC CTCCCATATG 3120
 GCACCGCCAT GGAAGAGCT CAGCTCAAGC CTCAGCCAC CTCTGATGCC TGAATCTCTC 3180
 CAAATTCAGA CTCCTATTCC TGAAGAACCA GTCCCCCACC CCTCATTTCT CTGAAAGGGA 3240
 35 GGGGTCTGGG TACTTCTGA AGGTGCTGAC GGCCAGGGAG AAGCTCTCTC CCCCAGCCCA 3300
 GAGACATACT TGAAGGCCA GAGCCAGGGG GGTGAGGAGC TGGGGATCCC TCCCCCCTAT 3360
 GCATGTGAA GGACCTTGT TTACACATAC CCTCTTCATG GATGGGGGAA CTCAGATCCA 3420
 GGGAGAGAG CCCAGCTCC CTGAAGCCTT TGCAATTGAG AGAGTTTCTC GAAACAACCTG 3480
 GAAAGATAAC TAGGAATCC ATTACAGTT CTTTGGGCCA GACATGCCAC AAGGACTTCC 3540
 40 TGTCCAGCTC CAACCTGCAA AGATCTGTCC TCAGCCTTGC CAGAGATCCA AAAGAAGCCC 3600
 CCAAGTAAGA CCTGGAACCT GGGAGTTTAA GACCTGGCAG CTCTGGACAG CCCCACCTG 3660
 GTGGGCCAAC AAAGAAGCT AACTATGCAT GGTGCCCCAG GACCAGCTCA GGACAGATGC 3720
 CACAGGATA GATGCTGGCC CAGGGCCAGA GCCCAGCTCC AAGGGGAATC AGAATCAAA 3780
 TGGGCCAGGA TCCAGCTGG GGTCTGAGT TGAATCGGAA CCCAGACTCA GACATTGGCA 3840
 CCAATCCAGG CAGATCCAGG ACTATATTG GGCCTGCTCC AGACCTGATC CTGGAGGGCC 3900
 45 AGTTCACCTT GATTAGGAG AAGCCAGGAA TTTCCAGGA CCTGAAGGGG CCATGATGCC 3960
 AACAGATCTG GAACCTCAGC CTGGCCAGAC ACAGGCCCTC CTGTTTCCCC AGAGAAAGGG 4020
 GAGCCACCTG TCTGGGCTC GCAGAAATTG GGTCTGCTC GCCAGCTGCA CTGATGCTGC 4080
 CCCTCATCTC TCTGCCAAC CCTTCCCTCA CCTTGGCACC AGACACCCAG GACTTATTTA 4140
 AACTCTGTGT CAAGTGCAAT AAATCTGACC CAGTGCCCCC ACTGACCAGA ACTAGAAAAA 4200
 AAAA

Seq ID NO: 6 Protein sequence:
 Protein Accession #: NP_002196.1

55 1 11 21 31 41 51
 MGRTPESPL HAVQLRNGPR RRPPLVPLLL LLVPPPPRVG GFNLDAEAPA VLSGPPGSPFF 60
 GFSVBFYRPG TDGVSVLVGA PKANTSQPGV LQGGAVYLCF NGASPTQCTP IEFDSKGSRL 120
 60 LESSLSSESG REPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPVGTCTYLS 180
 DNFTRILEYA PCRSDFSWAA GQGYCQGGF8 AEPTKTGRVV LGGPGSYFWQ GQILSATQEQ 240
 IAE5YYPEYL INLVQQLQT RQASSIYDDS YLGYSVAVGE FSGDDTEDFPV AGVPKGNLTY 300
 GYVITLNGSD IRSLYNFSGE QMASYFGYAV AATDVNGDGL DDLLVGAFL MDRTPDGRPQ 360
 EVGRVYVYLG HPAGIEPTFT LTLTGHDFFG RFGSSLTPLG DLDDQGVNDV AIGAPFPGST 420
 65 QQGVVVFVFG GPGGLGSKPS QVLQPLMAAS HTPDFPGSAL RGGRDLDNGV YPDLIVGSPG 480
 VDKAVVYRGR PIVSASASLT IPPAMFNPEE RSCSLBGNPV ACINLSPLCN ASGKHVADSI 540
 GFTVELQLDW QKQGGVRRR LFLASRQATL TQTLLIQNGA REDCREMKIY LRNESEFRDK 600
 LSPHIALNFP SLDPAQPVDS HGLRPALHYQ SKSRIEDKAQ ILLDCGEDNI CVPDLQLEVF 660
 GEQNHVYLGD KNAIINLTPEA QNVGEGGAYE AELRVTAFFE AEYSGLVRHP GNPSSLSCDY 720
 70 FAVNQSRLLV CDLGNPMKAG ASLWGLRFT VFHLRDTKKT IQDFQILSK NLANSQSDVV 780
 SFRLSVBAQA QVTLNGVSKP EAVLFPVSDW HPRDQPKKEE DLGPAVHEVY ELINQGPSSI 840
 SQGVLELSCP QALBQQQLLY VTRVTGLNCT TNHPIINPKGL ELDPGSLHH QKREAPERS 900
 SASSGPQLLK CPEACFRLR CELGPLHQBE SQSLQLHPRV WAKTFLQREH QPFSLQCEAV 960
 YKALMFPVRI LPRQLPQKER QVATAVQWTK AEGSYGVPLW IITLAILFGL LLLGLLIYIL 1020
 75 YKLGFPPKRLS FYGTAMEKAQ LKPPATSDA

Seq ID NO: 7 Nucleotide sequence:

Nucleic Acid Accession #: NM_002211

Coding sequence: 104..2500 (underlined sequences correspond to start and stop codons)

5
10
15
20
25
30
35
40
45
50
55
60
65
70

```

1      11      21      31      41      51
|      |      |      |      |      |
GTCGCGCAAA ACCTGCGCGG ATAGGGAAGA ACAGCACCCC GGCGCGGATT GCGGTACCAA 60
ACAAAGCCTAA CGTCCGCTGG GCCCGGAGCG CGCGCGGAA AAGATGAATT TACAACCAAT 120
TTTCTGGATT GGACTGATCA GTTCAGTTTG CTGTGTGTTT GCTCAACAG ATGAAATAG 180
ATGTTTAAAA GCAATGCGCA AATCATGTGG AGAATGTATA CAAGCAGGCG CAAATGTGG 240
GTGCTGCACA AATTCAACAT TTTTACAGGA AGAATGCGT ACTTCTGCAC GATGTGATGA 300
TTTAGAAGCG TTAATAAAGA AGGTTGCCCC TCCAGATGAC ATAGAAAATC CCAGAGGCTC 360
CAAAGATATA AAGAAAAATA AAAATGTAA CAAACGTAGC AAAGGAACAG CAGAGAGCT 420
CAAGCCAGAG GATATTACTC AGATCCAACC ACAGCAGTTG GTTTGCGAT TAAGATCAG 480
GGAGCCACAG ACATTTCAT TAAATTCAA GAGAGCTGAA GACTATCCCA TTGACCTCTA 540
CTACCTTATG GACCTGTCTT ATTCAATGAA AGACGATTGG GAGAATGTAA AAAGTCTTGG 600
AACAGATCTG ATGAATGAAA TGAGGAGGAT TACTTGGGAC TTCAGAAATTG GATTGGCTC 660
ATTTGTTGAA AAGACTGTGA TGCCCTACAT TAGCACAACA CCAGCTAAGC TCAGGAACCC 720
TTGCACAGAT GAACAGAAGT GCACACCCCC ATTAGCTAC AAAATGTGTC TCAGTCTTAC 780
TAATAAAGGA GAAGTATTTA ATGAAGTTGT TGAATAACAG CGCATATCTG GAAATTTGGA 840
TTCTCCAGAA GGTGGTTTGG ATGCCATCAT GCAAGTTGCA GTTTGTGGAT CACTGATTGG 900
CTGGAGGAAT GTTACACGGC TGCTGGTGTG TTCCACAGAT GCGCGGTTTC ACTTGTCTGG 960
AGATGGGAAA CTGGTGGCA TTGTTTACC AAATGATGGA CAATGTCACC TGAATAATA 1020
TAGTGTACAA ATGAGCCATT ATTAGATTA TCCTCTATT GCTCACCTTG TCCAGAACT 1080
GAGTGAATAA AATATTGAGA CAATTTTGC AGTACTGAA GAATTTGAG CTGTTTACAA 1140
GGAGCTGAAA AACTTGATCC CTAAGTCAGC AGTAGGAACA TTATCTGCAA ATTCTAGCAA 1200
TGTAATTTCG TTGATCATTT ATGCATACAA TTCCCTTCC TCAGAACTCA TTTTGGAAAA 1260
CGGCAATTTG TCAGAAAGAG TAACAATAAG TTACAATCT TACTGCAAGA ACGGGGTGAA 1320
TGGAACAGGG GAAATGGAA GAAATGTTT CAATATTTCC ATTGGAGATG AGGTTCAATT 1380
TGAATTAGAG ATAACTTCAA ATAAAGTATC AAAAAGGAT TCTGACAGCT TTAATAATTAG 1440
GCCTCTGGGC TTTCAGGAGG AAGTAGAGGT TATCTCTCAG TACATCTGTG AATGTGAATG 1500
CCAAAGCGAA GGCATCCCTG AAAGTCCCAA GTGTCAAGAA GGAATGGGA CATTGAGTG 1560
TGGCGGTGTC AGGTGCAATG AAGGGCTGTG TGTAGACAT TGTGAATGCA GCACAGATGA 1620
AGTTAAACAT GAAGACATGG ATGCTTACTG CAGGAAGAA AACAGTTGAG AAATCTGAG 1680
TAACAATGGA GAGTGCCTCT GCGGACAGTG TGTGTTAGG AAGAGGGATA ATACAAATGA 1740
AATTTATCTT GGCAAATCTT GCGAGTGTGA TAATTTCAAC TGTGATAGAT CCAATGGCTT 1800
AATTTGTGGA GGAATGGTG TTTGCAAGTG TCGTGTGTGT GAGTGCAACC CCACTACAC 1860
TGGCAATGCA TGTGACTGTT CTTTGTATAC TAGTACTTGT GAAGCAGCA ACGGACAGAT 1920
CTGCAATGCG CGGGGCTATC GCGAGTGTGG TGTCTGTAG TGTACAGATC CGAAGTTTCA 1980
AGGCGAAGCG TGTGAGATGT GTCAGACCTG CCTTGGTGTG TGTGCTGAGC ATAAAGAAATG 2040
TGTTCAGTGC AGAGCCTTCA ATAAAGGAGA AAAGAAAGAC ACATGCACAC AGGAATGTTTC 2100
CTATTTTAAAC ATTACCAAGG TAGAAGATCG GGACAAATTA CCCCAGCCGG TCCAACCTGA 2160
TCCTGTGTCC CATTTGTAAG AGAAGGATGT TGACGACTGT TGGTTCTATT TTACGTATTC 2220
AGTGAATGGG AACACGAGG TCATGTTTCA TGTGTTGGAG AATCCAGAGT GTCCCACTGG 2280
TCCAGACATC ATTCCAATTG TAGCTGGTGT GGTGCTGGA ATTGTTCTTA TTGGCCTTGC 2340
ATTACTGCTG ATATGGAAGC TTTTAATGAT AATTCATGAC AGAAGGGAGT TTGCTAAATT 2400
TGAAAGAGAG AAAATGAATG CCAATGGGA CACGGGTGAA AATCCTATTT ATAAGATGTC 2460
CGTAACAATC GTGGTCAATC CGAAGTATGA GGGAAAAAGA GTACTGCCCG TGCAAAATCCC 2520
ACAACACTGA ATGCAAGATA GCAATTTCCA TAGTCACAGT TAGGTAGCTT TAGGGCAATA 2580
TTGCCATGCT TTTACTCATG TGCAGGTTT TGAATGTAT AATTTTAAAA 2640
TGTTTTATTA TTTTGAATAT AATGTTGTAA TTCAATGCCG GAGCTGACAA AAGACTTGAG 2700
ACAGGATGCT TATTCTGTGC AGCTAAGGTC ACATTGTGCC TTTTGGACCT TTTCTTCTG 2760
GACTATTGAA ATCAAGCTTA TTGATTAAAG TGATATTCTT ATAGCGATTG AAAGGGCAAT 2820
AGTTAAAGTA ATGAGCATGA TGAGAGTTTC TGTTAATCAT GTATTAAAC TGATTTTATG 2880
CTTTACATAT GTCAGTTTGC AGTTATGCGA AATCCAAAGT AATGTCCTG CTAGCTAGTT 2940
AAGGATTGTT TTAATCTGTT TATTTTGTCT TTTGCTGTT AGACATGACT GATGACATAT 3000
CTGAAAGACA AGTATGTTGA GAGTTGCTGG TGTAAATAC GTTTGAAATA GTTGATCTAC 3060
AAAGGCCATG GGAATAATTC AGAGAGTTAG GAAGGAAAAA CCAATAGCTT TAAACCTGTT 3120
GTGCCATTTT AAGAGTTACT TAATGTTTGG TAACTTTAT GCCTTCACCT TACAAATCCA 3180
AGCCTTAGAT AAAAGAACCG AGCAATTTTC TGCTAAAAAG TCCTTGATTT AGCACTATTT 3240
ACATACAGGC CATACTTTAC AAAGTATTTC CTGAATGGGG ACCTTTTGAG TTGAATTTAT 3300
TTTATTATTT TTAATTTGTT TAAATCTGCG TGCTTCTAT CACCTCTTCT AATCTTTTAA 3360
TGTATTGTTT TGCAATTTTG GGTAAAGACT TTTTATGAG TACTTTTCT TTGAAGTTT 3420
AGCGGTCAAT TGGCCTTTT AATGAACATG TGAAGTTATA CTGTGGCTAT GCAACAGCTC 3480
TCACCTACGC GAGTCTTACT TTGAGTTAGT GCCATACAG ACCACTGTAT GTTTACTTCT 3540
CACCATTGTA GTTGCCCATC TTGTTTACA CTAGTCACAT TCTGTGTTTA AGTGCTTTTA 3600
GTTTTAACAG TTCA

```

Seq ID NO: 8 Protein sequence:

Protein Accession #: NP_002202

75
1 11 21 31 41 51
| | | | | |
MNLQPIFWIG LISSVCCVFA QTDENRCLKA NAKSCGECIQ AGPNCGWCTN STFLQEGMPT 60

5 SARCDLLEAL KKKGCPDDI ENPRGSKDIK KKKHVTNRK GTAKLKPED ITQIQPQLV 120
 LRLRSRGPQT FTLKFKRAED YPIDLYLMD LSYSMKDDLE NVKSLGTDLN NEMRRITSDF 180
 RIGFGSPVEK TVMFIITTP AKLRNPCTSE QNCTSPFSYK NVLSLTNKGH VPNEVLVGQR 240
 ISGNLDSPEK GFDALMQAV CGSLIGWRNV TRLLVPSIDA GFHAGDGKL GGIVLENDQ 300
 CHLENNMYTM SHYDYPISA HLKQKLEBN IQTIPAVTEB PQPVYKELN LIPKSAVOTL 360
 SANSNNVQL IIDAYNSLSS EVILENGKLS EGVTSYKSY CKNGVNGTGE NGRKCSNISI 420
 GDEVQPRI8I TSNKCPKDS DSPKIRPLGF TEEVEVILQY ICECEQSBG IPSPKCHEG 480
 NGTFEGACR CNEGRVGRHC ECSTDEVNSE DMDAYCRKEN SSEICSNHG CVCGQCVCRK 540
 10 RDNTEIYSG KPCBCDFNC DRSNGLIOGG NGVCKCRVCE CNPNYTGSAE DCSLDSTCE 600
 ASNGQICNGR GICRCGVCKC TDPKFGQQT CMCQTCLGVC AEHKECVQCR APNKGKKOT 660
 CTQCSYPMI TKVSRDKLP QPVQDPVSH CKKDVDDCW FYFTYSVNGN NEVMHVVEN 720
 PECPTGDDII PIVAGVAGI VLIQLALLI WKLMIHHR REPAPKPKK MNAKMDTGEN 780
 PIYKSAVTV VNPKYEGK

15 Seq ID NO: 9 Nucleotide sequence:
 Nucleic Acid Accession #: NM_002425
 Coding sequence: 23..1453 (underlined sequences correspond to start and stop codons)

20 1 11 21 31 41 51
 AAAGAAGTGA AGGCAGTGA GAATGATGCA TCTTGCAATC CTGTGCTGT TGTGTCTGCC 60
 AGTCTGCTCT GCTTATCTCT TGAGTGGGGC AGCAAAAGAG GAGGACTCCA ACAAGGATCT 120
 TGCCCGAGCA TACCTAGAAA AGTACTACAA CCTCGAAAAG GATGTGAAC AGTTTAGAAG 180
 25 AAAGGACAGT AATCTCATTG TTAATAAAT CCAAGGAATG CAGAAGTTC TTGGGTGGGA 240
 GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGGAGT 300
 TCCTGACGTT GGTCACTTCA GTCCTTTTCC TGGCATGCCG AAGTGGAGGA AAACCCACCT 360
 TACATACAGG ATTGTGAATT ATACACCAGA TTAGCCAAGA GATGCTGTTG ATCTGCGCAT 420
 30 TAGAGAAGCT CTGAAGTCT GGAAGAGGT GACTCCACTC ACATCTCCA GGCTGTATGA 480
 AGGAGAGGCT GATATAATGA TCTCTTCCG AGTTAAAGAA CATGGAGACT TTTACTCTTT 540
 TGATGCCCCA GGACACAGTT TGGCTCATGC CTACCCACTC GGACCTGGGC TTTATGGAGA 600
 TATTCTCTTT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATCTCT 660
 CGTTGCTGCT CATGAACCTG GCCACTCCCT GGGGCTCTTT CACTCAGCCA ACCTGGAAGC 720
 TTTGATGTAC CCACCTCTCA ACTCATTCAC AGAGCTCGCC CAGTTCCGCC TTTGCAAGA 780
 35 TGATGTGAAT GGCATTCACT CTCTCTACGG ACCTCCCTCT GCCTCTACTG AGGAACCCCT 840
 GGTGCCACA AAATCTGTTT CTTCGGGATC TGAGATGCCA GCCAAGTGTG ATCCTGCTTT 900
 GTCTCTGATG GCCATCAGCA CTCTGAGGGG AGAATATCTG TTCTTTAAAG ACAGATATT 960
 TTGGCGAAGA TCCACTGGA ACCCTGAACC TGAATTTTAT TTGATTTCTG CATTTTGGCC 1020
 40 CTCTCTTCCA TCATATTGG ATGCTGCATA TGAAGTTAAC AGCAGGACA CCGTTTAT 1080
 TTTTAAAGGA AATGAGTTCT GGGCCATCAG AGGAAATGAG GTACAGCAG GTTATCCAAG 1140
 AGGCAATCAAT ACCCTGGGTT TTCTTCCAAC CATAGAGAAA ATTGATGAGC CTGTTTCTGA 1200
 CAAGGAAAGG AAGAAAACAT ACTTCTTTGC AGCGGACAAA TACTGGAGAT TTGATGAAAA 1260
 TAGCCAGTCC ATGAGCAAG GCTTCCCTAG ACTAATAGCT GATGACTTTC CAGGAGTTGA 1320
 45 GCCTAAGGTT GATCTGTAT TACAGGCATT TGGATTTTTC TACTTCTTCA GTGATCATC 1380
 ACAGTTTGA TTTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAGA GTACACAGCTG 1440
 GTTACATGTC TAGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTAAAT AAATCTAATA 1500
 ATTATTCATC TAAATGATTA TGAGCCAAAA TGGTTAATT TTCTGCTCAT TTCTGTGACT 1560
 GAAGAAGATG AGCCTGTCAG ATATCTGCAT GTGTCATGAA GAATGTTTCT GGAATTTCTC 1620
 50 ACTTGCTTTT GAATGCACT GAACAGATT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680
 ATGATTTTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGGCCTGTTC 1740
 CTT

55 Seq ID NO: 10 Protein sequence:
 Protein Accession #: NP_002416

60 1 11 21 31 41 51
 MMHLAFVLV CLFVCSAYPL SGAAKEDSN KDLAQQLLEK YYNLEKDVQK FRKDSNLIV 60
 KKIQQMOKFL GLEVTKLDT DTLEVVRKPR CGVPDVGHFS SFGMPKWRK TLTLYRIVNY 120
 TPDLPDRAVD SAIEKALKVN EBTPLTFPR LYEGRADIMI SPAVKHGDG YSPDGPGLSL 180
 AHAYPPGPG YGDHFDDE KWTEDASGTN LFLVAAHELH HSLGLFHSAN TEALMYPLYN 240
 SFTBLAQRL SQDDVNGIQS LYGPFPASTE EPLVPTKSVF SGSEMPAKCD PALSPDAIST 300
 LRGEYLFKID RYFWRSHWN PEPEPELLISA EWPFLPSYLD AAYEVNSRDT VFLFKGNEFW 360
 65 AIRGNVQAG YPRGIHTLGF PPTIRKIDAA VSDKRRKTY FPAADKYWRP DENSSQMEQG 420
 FPLRIADDF GVEPKVDAVL QAPGPPYFPS GSSQFEFDPN ARMVTHILKS NSWLRC

70 Seq ID NO: 11 Nucleotide sequence:
 Nucleic Acid Accession #: XM_058189
 Coding sequence: 169..774 (underlined sequences correspond to start and stop codons)

75 1 11 21 31 41 51
 GAAGACAGCC TCAGCTCTTC AGTTGTTGAT CATTGTCTAT TGTCTCCAA ACAGTAAACC 60

AGTATTTCAC ACTGAGATTG TCGGCTGCGG GTATATTCCA ATTCCCGCTC TCCTCATGAA 120
 TATGAAGTGA AGGGCTCTGA CCCTGGAAGT GGTCTTAAGC AGGCAAAAT GGGGTCTCGG 180
 AAGTGTGGAG GCTGCTTAAG TGTGTTGCTG ATTCCGCTTG CACTTTGGAG TATAATCGTG 240
 AACATATTAT TGTATTTCCT GAATGGGCCA ACTTCCTATG CATCCAGCAA TAAACTCACC 300
 AACTAGCTGT GGTATTTTGA AGGAATCTGT TTCTCAGGCA TCATGATGCT TATAGTAACA 360
 ACAGTTCTTC TGGTACTGGA GAATAATAAC AACTATAAAT GTTGCCAGAG TGAAAACCTGC 420
 AGCAAAAAAT ATGTGACACT GCTGTCAATT ATCTTTTCTT CCCTCGGAAT TGCTTTTCTT 480
 GGATACTGCC TGGTCATCTC TGCTTTGGGT CTGTGCAAG GGCCATATG CCGCACCCCTT 540
 GATGGCTGGG AGTATGCTTT TGAAGGCACT GCTGGAAGTT TCCTTACAGA TTCTAGCATA 600
 TGGATTACCT GCCTGGAAAC TGCACATGTT GTGGAAGTGA ACATCATTTT ATTTTCCATT 660
 CTCATAACCC TCAGTGGGCT TCAAGTATC ATCTCCCTCA TCAGAGTAGT CATGCAACTA 720
 TCCAAGATAC TGTGTGGAAG CTATTCAATG ATCTCCAGC CTGGAATCAT TTGATAAGG 780
 ACAAATGTTT TTCCATTATC AAGACATGGC CATCTATCTA AATATTATAT CAACTGTGTA 840
 GACTTGAGGG CAATATTGAA ATGATGTGTC TTCTGCTATT TGGTGTATT TGTAAAAA 900
 TTTGCACTCC TCACTGCACA TGCAAGTATA CCACTCTTCC ATTTAGTATG TTTTAAAGT 960
 AATATGCATC AGAAACTTCA GAAATACTTC TGCCCTTTGA TCAAAACAAAT CCATTTCCAA 1020
 GAATCTGTAT TAGGGAAGTA AATAAGAATA TGAGAGAAAC CTTTATGCAA ATATGTATAT 1080
 TGCAACATTA TTTAATATTC TGGAAAAATT GAAACACCCC AAAATTCTAA ACTCAGAGGA 1140
 AGGATTAAAT AAGAGTGGT ACATACCTGT AATGTTTCTT GATATTAAAA AAAAAATTAA 1200
 ATAAAAATA AAGAGTACTA CATGTTGTGA AAA

Seq ID NO: 12 Protein sequence:
 Protein Accession #: XP_058189

25
 1 11 21 31 41 51
 MGSRKCGGCL SCLLIPLALW SIIVNILLYF PNGQTSYASS NKLTNYVWYF EGICPFGIMM 60
 LIVITVLLVL ENNNYKCCQ SEMSKKYVT LLSIIPSSLG IAFSGYCLVI SALGLVQGPY 120
 30 CRTLDGWBYA PESTAGRLT DSSIWIQCLE PAHVVEWNI LPSILITLSG LQVIICLIRV 180
 VMQLSKILCG SYSVIFQPGI I

35
 Seq ID NO: 13 Nucleotide sequence:
 Nucleic Acid Accession #: NM_005397
 Coding sequence: 251..1837 (underlined sequences correspond to start and stop codons)

40 1 11 21 31 41 51
 AAACGCCGCC CAGGAAGCAG CCGCCGCGCG CCGCGCTCCT CTGCCACTGG CTCTGCGGCC 60
 CAGCCCGGCT CTGCTGCAGC GGCAGGAGAG AAGAGCCGCC CGAGCGGAC TOGGGAGCCC 120
 CGGSCCAGAG CTTGCCCTCC GGAGCCACCC ACAGGCTTCC CCGGCGGGCG CCCACGCTCC 180
 TACCGCCCGG ACGCGCGGAT CCTCCGCGCG CACCGCAGCC ACCTGCTCCC GGCCAGAGAG 240
 45 CGACGACAGG ATGCGCTGCG CGCTGGCGCT CTGCGCGCTG CTGCTACTGT TGTCAACGCC 300
 CGCGCTGCTG CGCTGCTGCG CGCTGCGGTC GCGCTGCGCG TCGCCCTCCC AGAATGCAAC 360
 CCGACTACTC ACGGACTCAT CTAACAAAAA AGCACCGACT CCAGCATCCA GTGTACCAT 420
 CATGGCTACA GATACAGCCC AGCAGAGCAC AGTCCCACT TCCAGGCCA ACGAATCTT 480
 GGCTCGGTTC AAGCGACCA CCTTGTGTGT ATCCAGTAC TCAACGGGGA CTACAACCTT 540
 50 GGCTCAGCAA GTCTCAGGCC CAGTCAACAC TACGTGGCT AGAGGAGGCG GCTCAGGCAA 600
 CCTTACTACC ACCATCGAGA GCCCAAGAG CACAAAAAGT GCAGACACCA CTACAGTTGC 660
 AACCTCCACA GCCACAGCTA AACCTAACAC CACAAACGAG CAGAATGGAG CAGAAGATAC 720
 AACAACTCTT GGGGGGAAAA GCAGCCACAG TGTGACCACA GACCTCATAT CCACTAAGGC 780
 AGAACATCTG ACGACCCCTC ACCCTACAAG TCCACTTAGC CCGGACAAAC CCACCTTGAC 840
 GCATCTCTGT GCCACCCCAA CAACTCTGGG ACATGACCAT CTTATGAAAA TTTCAAGCAG 900
 55 TTTCAAGCACT GTGGCTATCC CTGGCTACAC CTTCAAGAG CCGGGGATGA CCACCAACCT 960
 ACCGTATGCG GTTATCTGCG AAAGAACTCA ACAGACTTCC AGTCAGATGC CAGCCAGCTC 1020
 TACGCGCCCTT TCCTCCAGG AGACAGTACA GCCCACGAGC CCGCAACCG CATTGAGAAC 1080
 ACTTACCCCT CCAGAGACCA TGAGCTCCAG CCCCACAGCA GCATCAACTA CCCACCGATA 1140
 60 CCCCAAAACA CCTTCTCCCA CTGTGCTCA TGAGAGTAAC TGGGCAAGT GTGAGGATCT 1200
 TGAGACACAG ACACAGAGTG AGAAGCAGCT CGTCTGAAC CTCACAGGAA ACACCTCTG 1260
 TGCAGGGGCG GCTTCGGATG AGAAATTGAT CTCACTGATA TGCCAGCAG TCAAGGCCAT 1320
 CTTCAACCCG GCCCAAGATA AGTGGCGCAT ACGGCTGGCA TCTGTTCCAG GAAGTCAGAC 1380
 CGTGTCTGTC AAAGAAATCA CTATTACAC TAAGCTCCCT GCCAAGGATG TGTACGAGCG 1440
 65 GCTGAAGGAC AAATGGGATG AACTAAAGGA GGCAGGGGTC AGTGACATGA AGCTAGGGGA 1500
 CCAGGGGCCA CCGAGAGAGG CCGAGGACCG CTTGAGCATG CCCTCATCA TCACCATCTT 1560
 CTGCATGGCG TCATTCTGCT TCCTCGTGGC GGCCTCTTAT GGCTGCTGCC ACCAGCGGCT 1620
 CTCCAGAGAG AAGGACGAGC AGCGGCTAAC AGAGGAGCTG CAGACAGTGG AGAATGGTTA 1680
 CCATGACAAAC CCAACACTGG AAGTATGGA GACCTCTCTT GAGATGCAGG AGAAGAAAGT 1740
 70 GGTCAACCTC AAGCGGAGC TGGGGGACAG CTGGATCGTC CCTCTGACA ACCTGACCAA 1800
 GGACGACCTG GATGAGGAGG AAGACACACA CCTCTAGTCC GGTCTGCGCG TGCGCTCCAG 1860
 CAGCACCAAG GAGCTCCAGA CCAACCAACC CAAGTGCGGT TTGATGGGG AAGGGAAGA 1920
 CTGGGGAGGG AGAGTGAACCT CCGAGGGGTG TCCCTCCCA ATCCCCAGG GGCCTTAAT 1980
 75 TTTCTCTTTT CAACCTGAAC AATCACATT CTGTCCAGAT TCCTCTGTA AATAACCCA 2040
 CTAGTGGCTG AGCTCAGTGC TGCTGGATGA TGAGGAGAT CAAGAAAAAG CCACGTAAGG 2100
 GACTTTATAG ATGAATAGT GGAATCCCTT CATTTGCGAG TGAGATTGCC GAGACCTGAA 2160
 GAGGGTAAGT GACTTGCCCA AGGTACAGAG CACTTGGTGA CAGAGCCAGG ATGAGAACAA 2220

5 AGATTCCATT TGCACCATGC CACACTGCTG TGTTCCATG TGCCCTTCCT CCAGAGCAGT 2280
 CCCGGGCAGG GGTGAAACTC CAGCAGGTGG CTGGGCTGGA AAGGAGGGCA GGGCTACATC 2340
 CTGGCTCGGT GGGATCTGAC GACCTGAAAG TCCAGCTCCC AAGTTTCTCT TCTCTACCC 2400
 CAGCCTCGTG TACCATCTT CCCACCTCT ATGTTCTTAC CCCTCCCTAC ACTCAGTGTT 2460
 10 TGTTCCCACT TACTCTGTTC TGGGGCTCTT GGGATTAGCA CAGGTATTTC ATAACCTTGA 2520
 ACCCTCTGTT CTGGATTGCG ATTTTCTCAC ATTTGCTTGG TGAGATGGGG GCTTAAACCA 2580
 CACAGTCTC CTGGGTGAA CCAGGTCTGC TTAGGGGACC TGCGTGCAGG TGAGGAGAGA 2640
 AGGGGACACT CGAGTCCAGG CTGGTATCTC AGGGCAGCTG ATGAGGSGTC AGCAGGAACA 2700
 CTGGCCCACT GCCCTGGCA CTCTTGCGAG AGGCCACCCA CGATCTTCTT TGGGCTTCCA 2760
 15 TTTCACCAAG GGAATAAAAT CTGCTGTAGC TAGTGAGAGC AGCGTGTTC TTTTGTGTT 2820
 CACTGCTCAG CTGATGGGAG TGATTCCTCG AGACCCAGTA TGAAAGAGCA GTGGCTGCAG 2880
 GAGAGGCCTT CCCGGGGCCC CCATCAGCG ATGTGTCTTC AGAGACAATC CATTAAGCA 2940
 GCCAGGAAGG ACAGGCTTTC CCTGTATAT CATAGGAAC TCAGGACAT TTCAAGTTGC 3000
 TGAGAGTTTT GTTATAGTTG TTTTCTAAC CAGCCCTCCA CTGCCAAGG CCAAAAGCTC 3060
 20 AGACAGTTGG CAGAGTCCA GTTAGCTCAT CTCATCTACT CTGATTCTCC TGTCACAGC 3120
 GAAAGAGAGG CTGGAAAGC GCAGTGCATG CTGGGTGCAT GAAGGGCAGC CTGGGGGACA 3180
 GACTGTTGTG GGAACGTCCC ACTGTCTTGG CTGGAGCTA GGCCTTGTCT TTCTCTTCT 3240
 CTGTGAGCCT AGTGGGCTCG CTGCGTCTCT CTTCAGTTT CTGCTGGCAT CTCAGGGGAA 3300
 CACAAAGACT ATGTCTATT CCCTATATAG GACTTTTATG GGCTCGGCG TTAGCTGCCA 3360
 25 TGTAAGAGGC TCCTAAGCAG TGGGCAATGT GAGGTTTCAT CTGATTGAGA AGGGGGAATC 3420
 CTGTGTGGAA TGTTGAACCT TCGCAATGCT CTCCATGTT CTGGGCTTAA ATTCCTGGG 3480
 ATCAGTAGG AAAATGGGCA GAATGCTTA GGGGAATGAA ATTGCCATT TTCCGGTGAA 3540
 ACGCCACACC TCCAGGCTCT TAAGAGTCAG GCTCCGCTG TAGTAGCTCT GATGAAATAG 3600
 30 GCTATCCACT CGGATGGCT TACTTTTAA AAGGGTAGGG GGAGGGGCTG GGAAGATCT 3660
 GTCTGCAACC ATCTGCTTAA TTCTTCTCT CAGTCTCTGA GCCATCTGAT ATCCTAGGGG 3720
 GAAAGAGGAG GCCAGGGGTT CACATAGGGC CCCAGCGAGT TTCCAGGAG TTAGAGGGAT 3780
 GCCAGGCTAA CAGTTCCAA AAACATCTGC CCCAGTCTC TAGTGTTCGG AGSTGGGCG 3840
 GATGAGAAC AGTCCCTGTT TGGGGGAAA CAGGAATCT TGTTAGGCTT GAGTGAGGTG 3900
 35 TTGTCTCTCT TCTGCCCCG CGCTGGGTT TCTCCACCA GTAGGTTTTC TGTGTGGTC 3960
 30 CCGTGGGAGA GGCAGACTG GATTATTCCT CCTTGTCTGA TCTTGGTCA CACTTACCA 4020
 GCCAGGCTT TTGACGGAGA CAGCAATAG GCCTCTGCA ATCAATCAA GGCCTCAACC 4080
 CTATGGCTCT TTGAGACAG ATGATGACT GCAAGGACTA GAGAGCAGGA GTGCCCTGGC 4140
 AGGTGGGTCC TGACTCTCT GACTCTCCAT CGCTCTGTC AAGGAGAAC CGGAGAGGCT 4200
 CTGGGCTGAT TCAGAGGTTA CTGCTTATA TTGCTCCAAA CTGTGTAGT CTAGGCTTAG 4260
 40 GACAGCTTCA GAATCTGACA CCTTGCCTTG CTCTTGCAC CAGGACACCT ATGTCAACAG 4320
 GCCAAGCAGC CATGCTCTA TAAAGGTCAT CATCTCTGC CACCTTTACT GGGTCTTAAA 4380
 TGCTCTCTGA TAATCTAGAG AGCATTGGGT CTGGGAAGAG GTAAGAGGAA CACTAGAAGC 4440
 TCAGCATGAC TTAACAGGTG TGTAGCAAG ACAGTTTATC ATCAACTCTT TCAGTGGTAA 4500
 45 ACTGTGGTGT CCCAAGCTG CACAGGAGGC CAGAAACCAC AAGTATGATG ACTAGGAAGC 4560
 40 CTACTGTCAT GAGAGTGGGG AGACAGGCG CAAAGCTTAT GAAGAGGTA CAGAAATATC 4620
 TTTGCGTGT AAGACAGAT ACGGGTTTAA TCTAGTCTAG GCRCCAGATT TTTTCCCGC 4680
 TTGATAAGGA AAGCTAGCAG AAGTTTATT TAAACCACTT CTGAGCTTT ATCTTTTGTG 4740
 ACAATATACT GGAGAACTT TGAAGACAA GTTCAAACTG ATACATATAC ACATATTTT 4800
 50 TTGATAAGT AAATCAGTG ACCATGTTAA CCTACCCCTG ACTGCTTTAA GTGAACATAC 4860
 45 TTTGAAAAG CATATAGTTA GCTGAGTAT GGGCAAGTTT TTTCTCTGGA CAGGAATGTA 4920
 AATGTCTTAC TGGAAATGAC AAGTTTGTG TTGATTTTT TTTTAAACA AAAAATGAAA 4980
 TATAACAGA CAACTTATG ATAAAGTATT TGCTGTGAT ATCAGGTGTT TGTGTTGTT 5040
 55 TTTTAAATT TAAATGCAA CCTGCCCCC TCCCAAGCAA AGTCACAGT CCAITTCAGT 5100
 50 AAGGTTGGA GTCAATATGC TCTGTTGGC AGGCAACCTT GTAGTCATGS AGAAAGGTAT 5160
 TCAAGATCT AGTCCAATCT TTTCTAGAG AAAAGATAA TCTGAAGCTC ACAAGATGA 5220
 AGTGACTTCC TCAAAATCAC ATGGTTGAG ACAGAAACA GATTAAACC TGGATCCACA 5280
 GACTGTGCGC CTCAGAAGGA ATAATCGGTA AATTAAGAA TGTACTCGA AGGTGCCAGA 5340
 ATGACACAAA GGACAGAAAT CCTTCCCAAG TTGTTACCT AGCAAGGCTA GGGAGGCAT 5400
 60 GAACACAAAC ATAAGAACTG GTCTTCTCAC ACTTCTCTG AATCAITTAG GTTTAAGATG 5460
 55 TAAGTGAACA ATTCTTTCT TCTGCCAAGA AACAAAGTT TGGATGAGCT TTTATATATG 5520
 GAATCTACT CAACAGGACT GAGGGACCAA GGAACATGA TGGGGAGGC AAGAGAGGGC 5580
 AAGAGTAAA ACTGTAGCAT AGCTTTTGT ACAGTCACTA GCTGATCCCT CAGGTCTGCT 5640
 GCAACACAG CATGAGGAC ACAGATGACT CTTTGTGTT GGTCTTTTG TCTGCAATGA 5700
 65 ATGTTCAACA GTTTGCCCG GAACTGGGG ATCATATATG TCTTAGTGA CAGGGGCTG 5760
 AAGTACACT GAATTTACTG AGAACTTGT TTGTAAGAAC TATAGTTAAT AATTATTGCA 5820
 TTTCTTACA AAAATATATT TTGGAATAAT GTATCTGTC AATTAAAGT

Seq ID NO: 14 Protein sequence:

Protein Accession #: NP_005388

65 1 11 21 31 41 51
 | | | | |
 MRCALALSAL LLLSTPPLL PSSPSPSPSP SPSQNTATYT TDSSNKTAFT PASSVTIMAT 60
 DTAQGSVTPT SKANEILASV KATTLGVSSD SPGTTTLAQQ VSGPVNTTVA RGGGSGNPTT 120
 70 TIESPKSTKS ADTTTATST ATAKENTTSS QNGAEDTNS GQKSSHSVTT DLTSTKAHL 180
 TTHPTTSPLS PRQPTLTHFV ATPTSSGHDH LMKISSSST VALPGYTFPS PGMITTLFSS 240
 VIQRTQOTS SQMPASSTAP SSQETVQPTS PATALRPTPL PETMSSSPTA ASTTHRYPKT 300
 PSPVAHESN WAKCBLETO TQSEKQLVLN LTGNLTCAGG ASDEKLISLI CRAVKATFNP 360
 75 AQKCGRLA SVPGSQTVVV KEITHTKLP AKDVYERLKD KWDELKEAGV SDMKLGDQGP 420
 PEZAEKRFNM PLIITIVCMA SPFLLVAAAL GCCHQLRSQR KDQQLTREL QTVENGYHDN 480
 PTLVEMTSS EMQEKVVSL NGELGDSWIV PLDNLTKDDL DEEBDTHL

Seq ID NO: 15 Nucleotide sequence:

Nucleic Acid Accession #: NM_004105

Coding sequence: 150..1631 (underlined sequences correspond to start and stop codons)

5
10
15
20
25
30
35
40
45
50
55

1	11	21	31	41	51	
CTAGTATTCT	ACTAGAACTG	GAAGATTGCT	CTCGAGTTT	TTTTTTGTT	ATTTTGTAA	60
AAAATAAAAA	GCTTGAGCAG	CAATTTCATAT	TACTGTGACA	GGTATTTTGT	CTGTGCTGTG	120
CRAGGTAACT	CTGCTAGCTA	AGATTTCACA	<u>TGTTGAAAGC</u>	CCTTTTCCTA	ACTATGCTGA	180
CTCTGGCGCT	GGTCAAGTCA	CAGGACACCG	AAGAAACCAT	CAOGTACACG	CAATGCACTG	240
ACGGATATGA	GTGGGATCCT	GTGAGACAGC	AATGCAAAAG	TATTGATGAA	TGTGACATTG	300
TCGCCAGACG	TGTAAAGGT	GGAATGAAAT	GTGTCAACCA	CTATGGAGGA	TACCTCTGCC	360
TTCCGAAAAA	AGCCCAAGAT	ATTGTCAATA	ATGAACAGCC	TCAGCAGGAA	ACACAACCA	420
CAGAAAGAAC	CTCAGGGGCA	ACCACCGGGG	TTGTAGCTGC	CAGCAGCATG	GCAACCACTG	480
AGATGTTGCG	CGGGGGTGGT	TTTGTGGCCA	GTGCTGCTGC	AGTCGCAGGC	CCTGAAATGC	540
AGACTGGCGG	AAATAACTTT	GTCACTCCGG	GGAACCCAGC	TGACCCCTAG	CGCATTCCCT	600
CCAACCTCTC	CCACCGTATC	CAGTGTGCAG	CAGGCTACGA	GCAAAAGTAA	CACAACGTGT	660
GCCAAGACAT	AGACGAGTGC	ACTGCAGGGA	CGCACAACCT	TAGAGCAGAC	CAAGTGTGCA	720
TCATTTTACG	GGGATCCTTT	GCATGTCACT	GCCCTCCTGG	ATATCAGAAG	CGAGGGGAGC	780
AGTGGGTAGA	CATAGATGAA	TGTACCATCC	CTCCATATTC	CCACCAAGAA	TGCGTGAATA	840
CACCAGGCTC	ATTTTATGTC	CAGTGCAGTC	CTGGGTTTCA	ATTGGCAGCA	AACACTATA	900
CCTGCGTAGA	TATAAATGAA	TGTGATGCCA	GCAATCAATG	TGCTCAGCAG	TGCTACAACA	960
TCTTGTGGTC	ATTCACTCTG	CAGTGCATTC	AAGGATATGA	GCTAAGCAGT	GACAGGCTCA	1020
ACTGTGAGAA	CATTGATGAA	TGCAGAACTT	CAAGCTACCT	GTGTCAATAT	CAATGTGTCA	1080
ATGAACCTGG	GAAATTTCTA	TGTATGTGCC	CCAGGGGATA	CCAAGTGGTG	AGAAAGTAGAA	1140
CATGTCAAGA	TATAAATGAG	TGTGAGACCA	CAAAATGAATG	CGGGGAGGAT	GAAATGTGTT	1200
GGAATATATC	TGGCGGCTTC	CGTTGTTATC	CACGAAATCC	TTGTCAAGAT	CCCTACATTC	1260
TAAACACAGA	GAACCGATGT	GTTCGCCAG	TCTCAAATGC	CATGTGCCGA	GAACTGCCCC	1320
AGTCAATAGT	CTACAAATAC	ATGAGCATCC	GATCTGATAG	GTCTGTGCCA	TCAGACATCT	1380
TCCAGATACA	GGCCACAAC	ATTTATGCCA	ACACCATCAA	TACTTTTCGG	ATTAATCTGT	1440
GAAATGAAAA	TGGAGAGTTC	TACCTACGAC	AAACAAGTCC	TGTAAGTGCA	ATGCTTGTGC	1500
TGCTGAAGTC	ATTATCAGGA	CCAAGAGAAC	ATATCGTGA	CCTGGAGATG	CTGACAGTCA	1560
GCAGTATAGG	GACCTTCCGC	ACAAGCTCTG	TGTTAAGATT	GACAAATAAT	GTGGGGCCAT	1620
TTTCACTTTA	GTCTTTTCTA	AGAGTCAACC	ACAGGCATTT	AAGTCAGCCA	AAGAATATTG	1680
TTACTCTAAA	GCACTATTTT	ATTTATAGAT	ATATCTAGTG	CATCTACATC	TCTATACTGT	1740
ACACTCACCC	ATAACAAACA	ATTACACCAT	GGTATAAAGT	GGGCATTTAA	TATGTAAAGA	1800
TTCAAGATTT	GTCTTTTATA	CTATATGTAA	ATTAGACATT	AATCCACTAA	ACTGGTCTTC	1860
TTCAAGAGAG	CTAAGTATAC	ACTATCTGGT	GAAACTTGA	TTCTTTCCTA	TAAAGTGGG	1920
ACCAGCAAT	GATGATCTTC	TGTGTGCTTC	AAGGAACTT	ACTAGAGCTC	CACCTAACAGT	1980
CTCATAGAGA	GGCAGCCATC	ATAACCATTC	AATAGCATGC	AAGGGTAAGA	ATGAGTTTTT	2040
AACGTGCTTG	TAAAGAAATG	GAAAGGTCA	ATAAGATAT	ATTCTTTTAT	AAAATGGGGA	2100
TCTGCCATAT	TGTGTGTTGT	TTTATTTTTC	ATATCCAGCC	TAAAGGTGGT	TGTTTATTAT	2160
ATAGTAATAA	ATCATTGCTG	TACAACATGC	TGTTTCTGT	AGGGTATTTT	TAAATTTTGT	2220
AGAAATTTTA	GATTGTGAAT	ATTTTGTAAA	AAACAGTAAG	CAAAATTTTC	CAGAATTTCC	2280
AAATGAACCC	AGATACCCCC	TAGAAAATTA	TACTATTGAG	AAATCTATGG	GGAGGATATG	2340
AGAAATATAA	TTCTTCTTAA	ACCACATTGG	AAGTACCTGC	AAGAAGCAAA	CTCGGAAAT	2400
ATAATACAT	CCTGAATTC	AGGCATTAC	AAGATGCAGA	ACAAATGGA	TAAAGGTAT	2460
TTCACTGGAG	AAGTTTAAAT	TTCTAAGTAA	AATTTAAATC	CTAACACTTC	ACTAATTTAT	2520
AACATAAATT	TCTCATCTTC	GTACTTGATG	CTCAGAGAGG	AAGAAATGA	TGATGGTTTT	2580
TATTCCTGGC	ATCCAGAGTG	ACAGTGAAT	TAAACAAAT	ACCTCTCTAC	CCAATTTCTAT	2640
GGAATATTTT	ATACGTCTCC	TTGTTTAAAA	TCTGACTGCT	TTACTTTGAT	GTATCATATT	2700
TTTAAATAAA	AATAAATATT	CCTTTAGAA	ATCACTCTAA	AA		

Seq ID NO: 16 Protein sequence:

Protein Accession #: NP_004096

60
65
70

1	11	21	31	41	51	
MLKALFLTML	TLALVKSQDT	EETITYTQCT	DGYEWDVFRQ	QCKDIDECDI	VFDACKGGMK	60
CVNHYGGYLC	LPKTAQLIVN	NEQPQGETOP	ABGTSGATTG	VVAASSMATS	GVLPGGGFVA	120
GAAAVAGPEM	QTGRNPFVIR	RNPADPQRI	SNPSHRIQCA	AGYEQSEHNV	QODIDECTAG	180
THNCRADQVC	INLRGSPACQ	CPFGYKRG	QCVDIDECTI	PPYCHQRCVN	TPGSEFYCQCS	240
PGFLAANNY	TCVDINECDA	SNQCAQQCYN	ILGSPICQCN	QYSLSSDRLL	NCEDIDECRT	300
SSYLQYQYCV	NEPGKFCMC	PQGYQVVRSR	TQDINECET	TNECREDEMC	WNYHGGFRCY	360
PRNQCQDPYI	LTPENRCVCP	VSNAMCRELP	QSIYKYMSI	RSRDSVPSDI	PQIQATTIYA	420
NTINTFRIKS	GNENGEFYLK	QTSFVSAMLV	LVKSLSGPRE	HIVDLEMLTV	SSIGTFRTSS	480
VLRIITIVGP	FSF					

Seq ID NO: 17 Nucleotide sequence:

Nucleic Acid Accession #: NM_018894

Coding sequence: 27..1967 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
5 AAACATTCA ACAATTAAT GGGTGTAAAG AACTGGAAAA CCTGGACTCC TACCACATGC 60
AGATAAAACC AATAGAGTGC AGAATAAGAC TCAAGTCAAG TAAGTAACGT TAAACACCAT 120
AAAGACACAT GGCCTTCTTT GTGTACATGA CATGCATTCT CAACAATGCA CTGACGGATA 180
TGAGTGGGAT CCTGTGAGAC AGCAATGCAA AGATATTGAT GAATGTGACA TTGTCCGAGA 240
CGCTTGTAAA GGTGGAATGA AGTGTGTCAA CCCTATGGA GGATACCTCT GCCTTCCGAA 300
AACAGCCAGG ATTATTGTCA ATAATGAACA GCCTCAGCAG GAAACACAAC CAGCAGAAGG 360
AACTCAGGG GCAACCAACG GGGTTGTAGC TGCCAGCAGC ATGGCAACCA GTGGAGTGT 420
GCCCGGGGGT GGTTTGTGG CCAGTGCTGC TGCACTCGCA GGCCCTGAAA TGCAGACTGG 480
CCGAAATAAC TTTGTCTATCC GCGGGAACCC AGCTGACCCCT CAGCGCATTC CCTCCAAACC 540
TTCCACCGT ATCCAGTGTG CAGCAGGCTA CAGCAAAAGT GAACACAACG TGTGCCAAGA 600
CATAGACGAG TGCACTGCAG GACGCGACAA CTGTAGAGCA GACCAAGTGT GCATCAATTT 660
15 ACGGGGATCC TTGCTATGTC AGTGCCTCC TGGATATCAG AAGCGAGGGG AGCAGTGCCT 720
AGACATGATC GAATGTACCA TCCCTCCATA TTGCCACCAA AGATGCGTGA ATACACCGG 780
CTCATTTTAT TGCCAGTGCA GTCTTGGGT TCAATTGGCA GCAACAACT ATACCTGGCT 840
AGATATAAAT GAATGTGATG CCAGCAATCA ATGTGCTCAG CAGTGCTACA ACATTCCTGG 900
TTCAATCTAT TGTCACTGCA ATCAAGGATA TGAGCTAAGC AGTGACAGGC TCAACTGTGA 960
20 AGACATGAT GAATGCGAAA CCTCAAGCTA CCTGTGTCAA TATCAATGTG TCAATGAACC 1020
TGGGAAATCT TCATGTATGT GCCCCAGGG ATACCAAGTG GTGAGAAGTA GAACATGTCA 1080
AGATATAAAT GAGTGTGAGA CCACAAATGA ATGCCGGGAG GATGAAATGT GTTGGAAATTA 1140
TCATGGCGGC TTCCGTTGTT ATCCACGAAA TCCTTGTCAA GATCCCTACA TCTTAACACC 1200
AGAGAACCGA TGTGTTTGCC CAGTCTCAAA TGCCATGTGC CGAGAACTGC CCCAGTCAAT 1260
25 AGTCTACAAA TACATGAGCA TCCGATCTGA TAGGTCTGTG CCATCAGACA TCTTCCAGAT 1320
ACAGGCCACA ACTATTATG CCAACACCAT CAATACTTTT CGGATTAAAT CTGGAATGA 1380
AAATGAGAGG TTCTACCTAC GACAACCAAG TCCTGTAAGT GCAATGCTTG TGCTCGTGAA 1440
GTCAATATCA GGACCAAGAG AACATATCGT GGACCTGGAG ATGCTGACAG TCAGCAGTAT 1500
AGGAGCCTTC CCGACAAGCT CTGTGTTAAG ATTGACAATA ATAGTGGGGC CATTTTCATT 1560
30 TTAGTCTTTT CTAAGAGTCA ACCACAGGCA TTTAAGTCAG CCAAGAAATA TTGTTACCTT 1620
AAAGCACTAT TTTATTTATA GATATATCTA GTGCACTAC ATCTCTATAC TGTACACTCA 1680
CCCATACAAA ACAATTACAC CATGGTATAA AGTGGGCATT TAATATGTAA AGATTCAAG 1740
TTTGTCTTAT TACTATATG TAAATTAGAC ATTAATCCAC TAACTGGTC TTCTTCAAGA 1800
GAGCTAAGTA TACATATCT GGTGAAACTT GGATCTTTTC CTATAAAAGT GGGACCAAGC 1860
35 AATGATGATC TTCTGTGGTG CTTAAGGAAA CTTACTAGAG CTCCTAAC AGTCTCATAA 1920
GGAGGCGACC ATCATAACCA TTGAATAGCA TGCAAGGGTA AGAATGAGTT TTTAATGCT 1980
TTGTAAGAAA ATGGAAAAGG TCAATAAAGA TATATTTCT TAGAAAATGG GGATCTGCCA 2040
TATTTGTGTT GGTTTTAT TCTATATCCA GCCTAAAGGT GGTGTGTTAT TATATAGTAA 2100
40 TAAATCATTG CTGTACAACA TGCTGGTTTC TGTAGGGTAT TTTTAATTT GTGAGAAATT 2160
TTAGATTGTG AATATTTTGT AAAAACAGT AAGCAAAATT TTCCAGAAAT CCCAGAAATGA 2220
AAGCATACCC CCTTAGAAAA TTATACTATT GAGAAATCTA TGGGGAGGAT ATGAGAAAAAT 2280
AATTCCTTTC TAAACCAAT TGGAATGAC CTGAAGAAGC AAATCGGAA AATATAATAA 2340
CATCCCTGAA TTCAGGCATT CACAAGATGC AGAACAAAAT GGATAAAAGG TATTTCACTG 2400
45 GAGAAGTTT AATTCTAAG TAAATTTAA ATCTAACAC TTCCTAAT TATAACTAAA 2460
ATTTCTATC TTGCTACTTG ATGCTCAGAG AGGAAGAAAA TGATGATGTT TTTTATCTCT 2520
GGCATCCAGA GTGACAGTGA ACTTAAGCAA ATTACCTCC TACCAATTC TATGGAATAT 2580
TTTATACGTC TCCTTGTTTA AAATCTGACT GCTTTACTTT GATGTATCAT ATTTTATAAT 2640
AAAAATAAAT ATTCTTTAG AAGATCACTC TAAAA

50 Seq ID NO: 18 Protein sequence;
Protein Accession #: NP_061489.1

1 11 21 31 41 51
55 MHSQQCTDGY EWDVPRQCK DIDBCDIVPD ACKGGMKCVN HYGGYLCLPK TAQIIIVNNEQ 60
PQQBTQPAEG TSGATTGVVA ASSMATSGVL PGGGFVASAA AVAGPEMQTG RNNFVIRNRP 120
ADPQRIPSNP SHRIQCAAGY EQSEHNVQD IDECTAGTHN CRADQVCINL RGSFACQCPP 180
GYQKRSEQCV DIBECTIPPY CHQRVNTFG SFYQCQSPGF QLAANNVTCV DINECDASNQ 240
CAQKCNILG SFICQCNQGY BLSDDLNCB DIDECRTSSY LCQYQCVNEP GKFSQCMCPQG 300
60 YQVVRERTCQ DINBCETTNE CREDEMKNY HGGPRCYPRN PCQDPYILTP ENRCVCPVSN 360
AMCRELPQSI VYKYSIRSD RSVPSDIPQI QATTIYANTI NTFRIKSGNE NGEFYLRQTS 420
PVSAMLVIVK SLSGPREHIV DLEMLTVSSI GTFRTSSVLR LTIIVGPPSF

65 Seq ID NO: 19 Nucleotide sequence;
Nucleic Acid Accession #: NM_006500
Coding sequence: 27..1967 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
70 ACTTGGTCT CGCCCTCCGG CCAAGCATGG GGCCTCCAG GCTGGTCTGC GCCTTCTTGC 60
TCGCGCCCTG CTGCTGCTGT CCTCGCGTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
CGCTTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCCTCTCCC 180
AGTCCCAAGG CAACCTCAGC CATGTCGACT GGTTTTCTGT CCACAAGGAG AAGCGGAGCG 240
75 TCACTCTCCG TGTGCGCCAG GGCAGGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGCG 300
TCAGCTCCA GGACAGAGG GCTACTCTGG CCTGACTCA AGTACCCCC CAAGACGAGC 360

5 GCATCTTCTT TGCCAGGGC AAGCGCCCTC GGTCCAGGA GTACCGCATC CAGCTCCGG 420
 TCTACAAAGC TCCGGAGGAG CCAACATCC AGGTCAACCC CTGGGCAATC CCTGTGAACA 480
 GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGAGGAA CGGTACCCC ATTCCTCAG 540
 TCATCTGGTA CAAGAATGGC CGGCTCTGA AGGAGAGAA GAACCGGTC CACATTCAGT 600
 CGTCCAGAC TGTGGAGTC AGTGGTTTGT ACACCTTGA GATATTCTG AAGGCACAGC 660
 TGGTTAAAG AGACAAGAT GCCCAGTTTT ACTGTGAGCT CAATCAACGG CTGCCAGTG 720
 GGAACACAT GAAGGAGTCC AGGGAAGTCA CGTCCCTGT TTTCTACCC ACAGAAAAG 780
 TGTGGCTGGA AGTGGAGCCC GTGGGAATGC TGAAGGAAG GACCGGTG GAAATCAGST 840
 GTTGGCTGA TGGCAACCT CCACCACTC TCAGCATCAG CAAGCAGAAC CCCAGCACA 900
 10 GGGAGGCAGA GGAAGAGACA ACCAACGACA ACGGGTCTCT GGTCTGGAG CCGCCCGGA 960
 AGGAACACAG TGGGCGCTAT GAATGTGAG CCGTGAACCT GAACACCAT ATATCGCTGC 1020
 TGATGGAACC ACAGGAACCT CTGGTGAAC ATGTGTCTGA CTTCCGAGTG AGTCCCGCAG 1080
 CCGCTGAGAG ACAGGAAGGC AGCAGCTTCA CCTGACCTG TGAGGCAGAG AGTAGCCAGG 1140
 ACCTCGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAAGT CTTGGAAGG GGGCTCTGTC 1200
 15 TTCAGTTGCA TGACCTGAAA CCGGAGGCGAG GAGGCGGCTA TCGCTGGTG GCGTCTGTGC 1260
 CCAGCATACC CGGCTGAAC CGCACACAGC TGGTCAAGCT GGCATTTTT GGGCCCTCT 1320
 GGATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAGAGAA TATGTGTGT AATCTGTCT 1380
 GTGAAGGTC AGGGCACCC CGGCCACCA TCTCTGGA CTCAACGGC ACGGCAAGTG 1440
 20 AACAAAGACA AGATCCACAG CGAGTCTGGA GACCCCTGAA TGTCTCGTG ACCCGGAGC 1500
 TGTGGAGAC AGGTGTGAA TGCAACGGCT CCAACGACCT GGGCAAAAC ACCAGCATCC 1560
 TTTCTCTGGA GCTGTCAAT TTAACCAACC TCACACAGA CTCACACA ACCACTGGCC 1620
 TCAGCATCTC CACTGCCAGT CCTCATACCA GAGCCACAG CACTCCACA GAGAGAAGC 1680
 CGCCGAGAG GAGAGCCGG GCGTGTGTC TGTGTGTC GATTGTGTG ATCTGTGTC 1740
 25 TGGCGTGTCT GGGCGCTGTC CTCTATTTC TCTATAGAA GGGCAAGCT CCGTGCAGGC 1800
 GCTCAGGAAA GCAGGAGATC AGCTGCCCC CGTCTGTAA GACCGAAGT GTAGTTAAG 1860
 TTAAGTCAGA TAAGCTCCA GAAGAGATGG GCTCCTGCA GGCAGCAGC GGTGACAGA 1920
 GGGCTCCGG AGACCAAGGA GAGAAATACA TCGATCTGAG GCATAGCCC CGAATCACT 1980
 CAGCTCCCTT CCGTCCCTGG ACCATTCCCA GCTCCCTGCT CACTCTTCT TCAGCCAAAG 2040
 30 CCGTCCAAAG GACTAGAGAG AAGCCTCCTG CTCCCTCAC CTGCACACC CCTTTCAGAG 2100
 GGCACCTGGG TTAGGACCTG AGGACCTCAC TTGGCCCTGC AAGCGCTTT TCAGGGACCA 2160
 GTCCACACC ATCTCTCCA CGTTGAGTGA AGCTCATCCC AAGCAAGAG CCCAGTCTC 2220
 CCGAGCGGT AGGAGAGTTT CTTGCAGAAC GTGTTTTTC TTTACACACA TTAGGCTGT 2280
 AAATACCTGG CTCTGCCAG CAGCTGAGCT GGGTAGCCTC TCTGAGCTGG TTTCTGCC 2340
 35 CAAGGCTGG CTTCCACAT CAGGTGTCAC CACTGAAGTG AGGACACACC GAGGCCAGGC 2400
 GCTGTCTCAT GTTGAAGTGC GCTGTTTACA CCGCTCCGG AGAGCACCC AGCGCATCC 2460
 AGAAGCAGCT CAGTGTGTC TGCCACCAACC CTCTCTGCT CTTCTTCAA GTCTCTGTG 2520
 ACATTTTTTC TTTGGTACA AGCCAGGAAC TGGTGTGATT CCTTAAAGA TACGTGCCGG 2580
 GGCCAGGTGT GGTGCTCAC GCTGTGTAAT CCAGCATTT GGGAGGCCGA GCGGGCGGA 2640
 40 TACCAAGATC AGGACGAGAC CATCTGGCT AACACGGTGA AACCTGTCT CTACTAAAA 2700
 TACAAAAAAT AATTAGTAG GGTGTGTGT TGGCACTAT AGTCCAGCT ACTCGAAGG 2760
 CTGAAGCAGG AGAATGTAT GAATCCAGGA GGTGAGCTT CAGTGCAGC GAGACCTGC 2820
 CACTGCCTC CAGCTGGGC AACACAGCGA GACTCGTCT CGAGGAAAA AAAAGAAAG 2880
 ACGGTACTT GCGGTGAGGA AGCTGGGCGC TGTTTTGGG TTCAGGTGA TTAGCTCAA 2940
 45 TCCCGGTGT CACTGTCTCC CATAGCCCTC TTGATGATC ACGTAAAACT GAAAGGCAGC 3000
 GGGAGCAGA CAAAGATGAG GTCTACACTG TCCTTCATGG GGAATAAAGC TATGGTTATA 3060
 TTAGCACCAG ACTTCTACAA ACCAAGCTCA GGGCCCAAC CCTAGAAGG CCCAATGAG 3120
 AGAATGGTAT TTAGGATGG AAAACGGGGC CTGCTAGAG CTTGGGTGT GTGTGTCTG 3180
 CTGTGTGTAT GCATACATAT GTGTGTATAT ATGTTTTGT CAGTGTGTA AATTGCAAA 3240
 50 TGTGTTCTT TATATATGA TGTATATATA TATATGAAA TATATATATA TATGAAAAAT 3300
 AAAGCTAAT TGTCCAGAA AATCATACAT TGCTTTTTTA TTCTACATGG GTACCAAGG 3360
 AACCTGGGG CCGTGTAAAC TACAACCAA AGGCACACA AACCGTTTC AGTTGGCAGC 3420
 AGAGATCAGG GGTACCTCT GCTTCTGAGC AAATGGCTCA AGCTCTACCA GAGCAGACAG 3480
 CTACCTACT TTTAGCAGC AAAACGTCCC GTATGACGA GCACGAAGG CCGGCAAGG 3540
 55 TGTAGCAGG AGCTATGTC CTTCCTATCG TTTCCGTCA CTT

Seq ID NO: 20 Protein sequence:
 Protein Accession #: NP_006491

60 1 11 21 31 41 51
 MGLPRLVCAF LLAACCCCPR VAGVPGEAEQ PAPELVEVEV GSTALLKCGL SQSQGNLSHV 60
 DWFSVHKERR TLIFRVRRGQ GQSEPGYEYQ RLSLQDRGAT LALTQVTPQD BRIFLCQGRK 120
 65 PRSQEYRIQL RYKAPBEPN IQVNPLGIPV NSKEPEEVAT CVRNGVPYIP QVIWYKNGRP 180
 LKEBKRNVRHI QSSQTVESSG LYTLOSLIKA QLVKEDKDAQ FYCELNVRLP SGNHMKESRE 240
 VTVEVFPYTE KWLVEVPVG MLKEGDRVEI RCLADGNPPP HFSISKQNPB TREAEETTN 300
 DNGVLVLEPA RKEHSGRYEC QAWNLDTMIS LLSEPPQLLV NYVSDVRVSP AAPERQSSSS 360
 70 LTLTCEAESS QDLFPQNLRE ETDQVLERGP VLQLHDLKRE AGGYRCVAS VPSIPGLNRT 420
 QLVLKLAIFGP PMAFKPKRV NVKENMVLNL SCEASGHPRP TISWNVNGTA SEQDQDPQRV 480
 LSTLNVLVNT ELLETGVECT ASNDLGKNTS ILFLVLVNL TLTPDSNTT GLSTSTASPH 540
 TRANSTSTER KLPEPESRGV VIVAVIVCIL VLAVLGAVLY FLYKKGLPC RRSKQZEITL 600
 PPSRKTSLVV EVSKDKLPEE MGLLQSSGD KRAPGDQGEK YIDLKH

	1	11	21	31	41	51	
	GGGATATTGG	ATGTAGCAAGA	GGCTGGGAAG	CCATCACTTA	CCTTGCACGT	AGAAAGAAGA	60
	CAGAAGGCAG	TATGCACAGC	TTTCTCTCAC	TGCTGCTGCT	CGTGTTCCTGG	GGTGTGGTGT	120
10	CTCACAGCT	CCACAGGACT	CTGAGAACAC	ATGAGCAAGA	TGTGGACTTA	GTCCAGAAAT	180
	ACCTGGAAAA	CTCTGACACT	CTGAGAAATG	ATGGAGGCCA	AGTTTGAAGA	CGGAGAAATA	240
	GTGGCCCGAT	GGTTGAAAAA	TTGAAGCAAA	TGCAAGAAAT	CTTTGGGCTG	AAAGTGACTT	300
	GGAAACCCGA	TGCTGAAACG	GAGGAAGTTA	TGAGCAGACC	CAGATGTGGA	CTGCTCTGATG	360
	TGBCATCAGT	TGCTCTCACT	CAGAGGAGAC	TCGCTGTGGG	CAGAACCACT	CTGACCTTACA	420
15	GGATTGAAAA	TTACACGCCA	GATTTGCCAA	GAGCAGATGT	GGACACTGCC	ATTGAGAAAG	480
	CCTTTCCAAT	CTGGAGTAAT	GTACACACTC	TGACATTTCAC	CAGAGTCTCT	GAGGGTCAAG	540
	CAGACATCAT	GATATCTTTT	GTCAAGGGAG	ATCATCCGGG	CAACTCTCCT	TTTATGTGAC	600
	CTGGAGGAAA	TGCTGCTCAT	GCCTTTTCAAC	GACGGCCGAG	TATTTGGGAG	GATGCTCAAT	660
	TTGATGAAGA	TGGAAGATGG	ACCAACAATT	TCAGAGAGTA	CACTTACATC	CGTGTTCGGG	720
20	CTCATGAAT	CGGCCATTCT	TTGSGACTCT	CCCATTTCTAC	TGATATCGGG	CTTTTGATGT	780
	ACCCTAGCTA	CACCTTCACT	GGTGTATGTT	TGATAGCTCA	GGATGACATT	GATGGCATCC	840
	AAGCCATATA	TGCGCTTCTG	CAAAATCTCG	TCAGCCCATC	CGCCGCCAAA	AGCCCAAAG	900
	CGTGTGACAG	TAAAGCTAAC	TTTGATGCTA	TAACTACGAT	TCGGGGAGAA	GTGATGTTCT	960
25	TTAAAGACAG	ATTTCATACG	CGCACAATCT	CGCTTACACC	GGAAATGTAG	CTCAATTTCA	1020
	TTTCTTGTTT	GTGCGCCAAA	CTGCGAAATG	GGCTTGAAAG	TGTTACGAGA	TGTTCGGACA	1080
	GAGATGAAGT	CCCGTTTFTT	AAAGGCGAAT	AGTACTGGGC	TGTTTACGGA	CAGAATGTGC	1140
	TACAAGGATA	CCCGAAGGAT	ATCTGACAGT	CCCTTGCTCT	CTCCAGAAGT	TGTAAGACATA	1200
	TCGATGCTGC	CTCTTCTGAT	GAAGAACACT	GAARAACTCA	CTTCTTTGTT	GTCTAACAACT	1260
30	ACTGAGGAGT	TGATGAATAT	AAACAGTACT	TGGATCCAGG	TTATCCCAAA	ATGATAGTAGC	1320
	ATGACTTTCC	TGGAATTGCC	CAACAAGTTG	ATGACTGTTT	CATGAAGAAT	GGATTTTCTC	1380
	ATTTCTTTCA	TGGAACAGCA	CAATCAAAAT	TTCAGTCTAA	ACGAGAGAGA	ATTTTGACTG	1440
	TCCAGAAAGC	TAAATAGTGT	TTCAACTGCA	GGAAAAATTG	JACATATTCTA	ATTTGAATGT	1500
	AAACACATTA	GTGTAGCTGC	AAAGAAGTGC	TTTCTTCGAA	GACCTGTCTA	TTTTCTCAGT	1560
35	CATTTTTTAAC	CTCTAGAGTC	ACTGATACAC	AGAAATATAAT	CTTATTTTATA	CCTCAGTTTG	1620
	CATATTTTTT	TACTATTTAG	AATGTAGCCC	TTTGTGTACT	GATATATAATT	AGGTTCCACA	1680
	ATGGTGGGTT	CAAAAAGTCA	AGTTTGTGGC	TTATGGATCT	ATATAGGCGA	GAGTTGCGAA	1740
	GATCTTTTCC	AGAGATATGA	ACTTCGAGGT	TGATCCCGCA	GAGCAGCTTC	ATGTGACAACT	1800
	ATATCTCTTT	AGAGACAAGCA	GAGTCAAGCT	ACATGAGTCT	TTCGCCGAGT	AAAGCAGAGT	1860
	CAGAACAACA	TGTGCACTCA	CTGGTGTAA	CTCGGTATAG	CAAGGAGTAA	CTTCTCTAAC	1920
40	ACAAAAATAG	TGTTTTATGT	TGGAATATAC	GTCAACCTTG	TTTCTACTGT	TTT	

Protein Accession #: NP_002412

45

1 11 21 31 41 51

50 MHSFPPLLLL LFWGVVSHSP PATLETQEQD VDLVKQYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLLQMQQE PGLKVTGKPD AETLKVMKQP RCGVDPDVAQ VLTEGNFRWE QTHLYTIRIN 120
YTPDLPPADV DHAIEKAPQL WSNVTPLTFT KVSEQADIM ISFVGRGDHR QSFYDGGPGGN 180
LAHAFQPVAG IGGDAHFPED ERTWNFRFYE NLRHVAHEL HSGLSLHST DIGALMYPSY 240
TFSGVDVQLAQ DDIDGQIAIY GRSQNPVQPI GPQTAKADCS KLTFDAITTI RQGVMMFFKDR 300
55 PKDIYSSGFP EVELNFIQV WPQLPNGLBA AYSBPRACDE RFFKGNKYNA VQGQIVHLRG 360
GKDIYSSGFP PRTVKHIDLA LSEENTGKTY FFKANKYNYR DEYKRSMDPN YPKMIAHDFP 420
GIGHKVDAVE MKDGFFFYGH TROYKEKDPK TKRILITLOKA NSWFNCRKN

65	1	11	21	31	41	51	
	TCTGCGTG	CCGGGGCTAG	GGGCTGGAAG	TCCTGGCTCT	AGTTGACAC	CGGAAGGAAA	60
	AGGCAACAAC	AGGAGGGGAA	GGGCTTAGAG	ACTGCTCTGA	TCGACAGCAC	TTTCTCTGGC	120
	CTCTACAGG	CTGTGTGCT	<u>ATGGGTTCC</u>	CGCCGGCCG	GAGGGGAGGG	CTGGGCTATG	180
	TCGGCGATCT	CACTCGCCAC	TCCTTCGACG	TGTGTGGGCA	CCCTCAACGAG	CTCGGCTCTG	240
70	CGGGGATCT	CATCAGATCT	AGCGTGTCTG	TGGCGGGACA	ACCCCTCAGA	GCACCAACGG	300
	CAGTTCTCAT	CGCCTCGAAT	CGGCTTCTCT	AGTCAATTTT	CGCGGGCCGT	CGGGGAGTGT	360
	GGGTGACAGT	GCTCTCTCTG	GCGGCTTCTG	CCGAAGGGCA	AGGCTTCAGT	CCTCTATTGG	420
	CTCTCATGTA	CACCTTCGCG	CGGCGCTCT	CTCCAGCCAC	TGCACACAGA	GTCTTAGCGG	480
	CGCCCACTTA	TTTGACAGTG	GAGCAGATGG	TCACAGCATG	CGACCGCTTC	ATCCCTGACG	540
75	CGCTATGAAC	CTTGGGCATC	TCCTCTGGCC	CTCTGGGAAG	AGAACCCCCA	ACACCCCTGA	600
	GGCCCGCTTC	ACCAGGTAAT	CCAGGCGGCT	CCGAGGAGCA	CCGACGACCA	CTCTACTGAAT	660

CTCGAAGCTG CAGTCAAGGC CCCCCAGTC CAGCCAGCCC TGACCCCAAG GCCTGCAACT 720
 GGAAAAAGTA CAGTACATC GTGCTAAACT CTCAGGCTCT CCAAGCAGGG AGCCTGGTCG 780
 GGGAGAGAAG TTCTGGTCAA CTTTGCCCCC AAGCCAGGCT CCCCAGTGBA GACGAGGCCT 840
 CCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GTGAAGAAGG ACCCATTCCT GGTCCCCAGA 900
 GCAGGCTCTC TCCAACTGCT GCCACTGTGC AGTTCAAATG TGGGGCTCCA GCCAGTACCC 960
 CCTACCTCCT CACATCCCAG GCTCAAGACA CCTCTGGATC ACCCTCTGAA CGGGCTCGTC 1020
 CACTACCGGG AAGTGAATTT TTCAGCTGCC AGAACTGTGA GGCTGTGGCA GGGTCTCAT 1080
 CGGGGCTGGA CTCTTGCTT CTGCGGAGC AAGACAAACC CTATAAGTGT CAGCTGTGCC 1140
 GGTCTTGCTT CGCTACAAG GGCAACCTTG CCAGTCAATG TACAGTGAC ACAGGGGAAA 1200
 AGCCTTACCA CTGCTCAATC TGCGAGGCC GTTTTAACCG GOCAGCAAAC CTGAAAACGC 1260
 ACAGCGCAT CCATTGCGGA GAGAAGCCGT ATAAGTGTGA GAGTGTGGC TCGGCTTTG 1320
 TACAGGTGGC ACATCTGGG GCGCACGTGC TGATCCACAC CGGGAGAGA CCCTACCTTT 1380
 GCCCTACCTG CGGAACCCGC TTCCGCCACC TGCAGCCCT CAAGAGCCAC GTTCCATCC 1440
 ACACCGGAGA GAAGCCTTAC CACTGCGACC CCTGTGGCT GCATTTCGG CACAAGATC 1500
 AACTGCGCT GCATCTGCG CAGAACACG GAGCTGTAC CAACACCAA GTGCACTACC 1560
 ACATTCTCG GGGGCCCTAG CTGAGCGCAG GCCCAGGCC CACTGTCTC CTGCGGTGG 1620
 GAAAGCTGCA GGCCCAAGCC TTGCTTCCCT ATCAGGCTG GGCATAGGG TGTGCCAGC 1680
 CACTTTGTGA TCAGAAATG CCACCTCTT AATTCTCAC TGGGAGAGC AGGGGTGGCA 1740
 GATCTGCTG AGATCTGCT CTGTTTGTCT GGTCAAAACC TCTTCCAC AAGCCAGATT 1800
 GTTCTGAGG AGAGAGCTAG CTAGGGGCTG GGAAGGGGA GAGATTGGAG TCCTGGTCTC 1860
 CTTAAGGGA TAGCCCTCCA CCTGTGGCC CCATTGCATT CAGTTATCT GTAAATATA 1920
 TTTATGAGG CCTTGGGTG GCACCGGGC CTTCATGGA TTGCATTTC CACTCCCTC 1980
 TTCACAGT GTGATTAAA GTGACCAGAA ACACAGAAG TGAGATCACA GCTCTGCTG 2040
 CAGAGATTAC TAGCCCTTGG CTCTCTGTT TGGCTTGGT ATTTATATT ATTTCTGTC 2100
 TAACCTTAT CTTTGAATG GTTCTTCTC CTGTTTGTG GCTTGTAGT TTGTTAAAA 2160
 TGGAAAAAG GGTCTCTGT GTTCTGCCC TGTAAATCTA GGTCTGGAAC CTTTATTGT 2220
 TCTAGGCGC CTCTGGGAA ATGCGGGAT GTGGAATGG GTCAGGAAC CTCTCTGTA 2280
 TCTGGAATG TGTAGTTCT CTAGCAGTCT AGAAATGGAT ACAGACATT CTCTGTTCT 2340
 CAGGGTGTAT AGGAACCAT ATGTTGAGC CAAATGGAA GTAATATAA ATGCTCTCT 2400
 GAGGCTGTG GTGTGGGGA TTCTGTATCT GGAATCCCTA TCACTCCAAC TGGAGGCTG 2460
 GGTGTGGGG GATTCTGTAT CTGGATTCCG TATCACTCCA AGTGGAGCT AGCAGGTTT 2520
 TCTGCAAGAT GGTCCAGAA CTAAATGTCT CCATTATCT GGTCACTGG GTTTGGCTCT 2580
 GCTGTATCCA TCTATAGTG TAGAGACCA CCAGGCTCA AGTGGAGTCC ATCATCTCC 2640
 CACGGGGGCC TGTCTTAGC ACTGAGTGA TCGCTCCAT GGGGAGAGAT CAGACATTCC 2700
 TTATCAGAGA TGATGTGACC TTTCTGACT CTGCCAGTC TCTATGAATG TTATGGCTTA 2760
 GGGAGAATC ATGAACTCT TTAGCTTGT TAGATGTAA ACAGTGTAA CCACTCCTT 2820
 ACTACAGAG CATATGGGT TGAATGTAC CTGGGTTCT CTCTATTGAG TTGAGCCCT 2880
 TCTTCTTTA GTGGGTTTG GACATCTCT GGCAGTGT CAGATGCCAG AACCTTCTT 2940
 TCTCTAGAA GGGATGCTG TTGTAACCT TACCTTTAA AAGCTGGTC TGTGACCTG 3000
 TCTTCCCATC CTGCATTCC TGTCTGGAAC CAGTGAATG ATTAGAACC TCCATAGGA 3060
 AAGAAAGGG GCTGAGTTC ATTTCTGGT TGTGTAGTT TGGTGGGAT TATTGTGGC 3120
 ATTACAGAT TAAAGATTG ACTAGCCAT AGGCCAAGG CCGTTCTAG TTGACCAAG 3180
 TTCAGTAGG ATTAAGAGG TGGTTAGGG GTGCAATTC TGTGTAGGC CAGGTAGGA 3240
 GAAGGTAGA AACAGGTTG CCTCTTGGT GGGTGGAGT TCTGAAATG TAGAAGAAGC 3300
 GCTGAAGCT TGATTGATG TTCTGCCCT TGTGCCCTG GGGCTTATCT GATTATGGGA 3360
 CGAGGGTGA AAGTAAGAAG CACTTTTGA TTTGTGGGT AGAACTTCAA CAATAAGTCA 3420
 GTTCTAGTG CTGTGCTG GGGACTAGT AGAAGCTAC TCTTCTCCT CTTCCTCTT 3480
 TCTCCCATG GCCCACTGC AGAATTAAAG AAGGAAGAAG GGAAGCCGA GAGCTCTATA 3540
 AGAAGGAAT ATGATTCTA TTAGCAGAT TGGATGGCA GGTGGAGAT GCCTGGGGT 3600
 AGAAATGTA GATCTTGCA CATCAGATC TTGAATATA GAAGCTCTC TGYCWRAAA 3660
 AAAAAAAA AAAAAA

Seq ID NO: 24 Protein sequence:

Protein Accession #: FGENESH predicted

1 11 21 31 41 51
 60 MGSPAAPBGA LGYVREFTRH SSDVLNLNE LRLRGILTDV TLLVGGQPLR AHKAVLIACS 60
 GPFYSIFRGR AGVGVVDVLSL PGGPEARFPA PLLDFMYTSR LRLSPATAPA VLAATYLMQ 120
 EHVVQACHRF IQASYEPLGI SLRPLEAEPP TPPTAPPPGS PRSEGHDPD PTESRSCSQG 180
 PPSPASPDPK ACNWKYKYI VLNSQASQAG SLVGRSSSG PCPQARLPFG DEASSSSSSS 240
 SSSSEGPPI GPQSRLSPTA ATVQPKCGAP ASTPYLLTSQ AQDTSGSPSE RARPLPGSEF 300
 65 PSCQNECEVA GCSGLDSLVL PGDRDKPYKC QLGRSSPYRK GNLSHRTVH TGEKPYHCIS 360
 CGARFNRPAN LKTHSRHSQ EKPYKCTCG SRFVQVHLR AHVLIHTGEK FYPCTCOGR 420
 FRHLQLKSH VRIHTGEKPY HCDPCGLHFR HKSQRLHLR QKGAATNTK VHYHILGGP

Seq ID NO: 25 Nucleotide sequence:

Nucleic Acid Accession #: U21551

Coding sequence: 1..1155 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 75 ATGCAATTGCA GTAACGGATC GGCAGAGTGT ACCGGAGAAG GAGGATCAAA AGAGGTGGTG 60

5 GGGACTTTTAA AGGCTAAAGA CCTAATAGTC ACACCAGCTA CCATTTTAAA GGA AAAACCA 120
 GACCCCAATA ATCTGGTTTT TGGAACTGTG TTCACGGATC ATATGCTGAC GGTGGAGTGG 180
 TCCTCAGAGT TTGGATGGGA GAAACCTCAT ATCAAGCCTC TTCAGAACCT GTCAATTGCAC 240
 CCTGGCTCAT CAGCTTTGCA CTATGCAGTG GAATTATTGT AAGGATTGAA GGCATTTCGA 300
 GGAGTAGATA ATAAAAATCG ACTGTTTCAG CCAAACTCA ACATGGATAG AATGTATCGC 360
 TCTGCTGTGA GGGCAACTCT GCCGTATTT GACAAAGAAG AGCTCTTAGA GTGTATTCAA 420
 CAGCTTGTGA AATTGGATCA AGAATGGGTC CCATATTCAA CATCTGCTAG TCTGTATATT 480
 CGTCTGTGAT TCATTGGAAC TGAGCCTTCT CTGGAGTCA AGAAGCCTAC CAAAGCCCTG 540
 10 CTCTTTGTAT TCTTGAGCCC AGTGGGACCT TATTTTTCAA GTGGAACCTT TAATCCAGTG 600
 TCCTTGTGGG CCAATCCCAA GTATGTAAGA GCCTGGAAAG GTGGAACCTG GGAAGTCAAG 660
 ATGGAGAGGA ATTACGGCTC ATCTCTTTT GCCCAATGTG AAGACGTAGA TAATGGGTGT 720
 CAGCAGGTCC TGTGGCTCTA TGGCAGAGAC CATCAGATCA CTGAAGTGGG AACTATGAAT 780
 CTTTTCTTTT ACTGGATAAA TGAAGATGGA GAAGAAGAAC TGGCACTGCC TCCACTAGAT 840
 15 GGCATCATTC TTCCAGGAGT GACAAGGCGG TGCATTTCTG ACCTGGCACA TCAGTGGGGT 900
 GAATTTAAGG TGTCAAGAGG ATACCTCACC ATGGATGACT TGACAACAGC CTTGGAGGGG 960
 AACAGAGTGA GAGAGATGTT TAGCTCTGGT ACAGCCTGTG TTGTTTGGCC AGTTTCTGAT 1020
 ATACTGTACA AAGCGAGAC AATACACATT CCAACTATGG AGAATGGTCC TAAGCTGGCA 1080
 AGCCGCATCT TGAGCAAAAT AACTGATATC CAGTATGGAA GAGAAGAGAG CGACTGGACA 1140
 20 ATTGTGCTAT CCTGA

Seq ID NO: 26 Protein sequence:
 Protein Accession #: AAB08528

25 1 11 21 31 41 51
 MDCSNGSARC TEGGGSKEV GTFKAKDLIV TPATILKEKP DPNNLVFGTV FTDHMLTVEW 60
 SSEFGWEKPH IKPLQNLSLH PGSSALHYAV ELFEGLEKAFR GVDNKIRLPQ PNLNMDRMYR 120
 30 SAVRATLFPV DKESLLECIQ QLVKLDQEMV PYSTSASLYI RPAFIGTEPS LGVKKPTKAL 180
 LFWLLSPVGP YFSSGTFNEV SLWANPKYVR ANKSGTGDCK MGNVYSSLF AQCEDVDNGC 240
 QQVLMLYGRD HQITEVOTMN LFLYWINEDG EERLATPPLD GIILPGVTRR CILDLAHQWG 300
 EFKVSELYLT MDDLTTALEG NRUREMFSSG TACVVCVPSD ILYKGETIHI PTMENGPKLA 360
 SRILSKLTDI QYGRESDDWT IVLS

35 Seq ID NO: 27 Nucleotide sequence:
 Nucleic Acid Accession #: XM_039209
 Coding sequence: 656..2758 (underlined sequences correspond to start and stop codons)

40 1 11 21 31 41 51
 TCGCGCGGGG GCGGCCCCCT CCCCTTCCCT CCACCCCTGGG CGGGGGCGCG CGAGAAGCGG 60
 TGACCTCAAG GGGCGCGCTG TGGCAGCACC TCCCGCGCGG CTAGTTAAAA AGAAGAAGAA 120
 45 AAGAGGGGAA GAAACATGAG AGGCTGTGTG AAGAACTGCA GCCCGCGGCA GAGGAGACCT 180
 GACATCATC TAGAGCCAG CGCTGGCCCT GCCTCGCGCT GCCCGCGCGC CGCGCTGCGC 240
 GTTTCCTGTC CTGCTACTGT CCACTCTAAA CAACTCCCGT TACACGGACA AGTGAACATC 300
 TGTGGCTGTC CTCTCTTTT CTCTCTCTCT TCCCACTCC TCTCTCTCT CCACTCTCC 360
 50 AGCCGACGCA GAAAGCCCCC AACCCAACTG ACACCTGGC CAACTGCAAC GGTGTCTATC 420
 GCACAACTTT ATCTCGCTCC TCGGGCTCCC CTAAGGCATT GACCCCATCG CCGCGTCTTT 480
 TATTTTTCG AAGTTGTCAT CGCTGTACAT ATTTTGTGCC CGCGCACCTC CCTCTGTCTC 540
 TGGAGTGCCC TACAGCCCGC CAAACTCTCT CTGGAGCTGC GCCCTAGTGC CCTGCTGGG 600
 CAGTGGCGTT CCCCCCATC CTCCCGCGCC CAGCCCGTGC TGCTCTGGGC AGACGATGCT 660
 55 GAAGATGCTC TCCTTTAAGC TGCTGCTGCT GCGCGTGGCT CTGGGCTTCT TTGAAGGAGA 720
 TGCTAAGTTT GGGGAAAGAA ACGAAGGGAG GGGAGCAAGG AGGAGAAGGT GCCTGAATGG 780
 GAACCCCGCG AAGCGCTGTA AAGGAGAGA CAGGAGGATG ATGTCCAGC TGGAGCTGCT 840
 GAGTGGGGGA GAGATGCTGT GCGGTGGCTT CTACCTCGG CTGTCTGCT CCGTGGGAG 900
 TGACAGCCCG GGGCTAGGGC GCCTGGAGAA TAAGATATT TCTGTTACCA ACAACACAGA 960
 60 ATGTGGGAAG TTACTGGAGG AATCAAAATG TGCACTTTGC TCTCCACATT CTCAAAGCCT 1020
 GTTCCACTCA CTTGAGAGAG AAGTCTTGA AAGAGACCTA GTACTTCTC TGCTCTGCAA 1080
 AGACTATTGC AAGAATTTCT TTTACACTTG CCGAGGCCAT ATTCAGGTT TCTTCAAAAC 1140
 AACTGCGGAT GAGTTTGTCT TTTACTATGC AAGAAAAGAT GGTGGGTGT GCTTTCCAGA 1200
 TTTTCCAAGA AAACAAGTCA GAGGACCAAG ATCTAACTAC TTGACCCAGA TGGAAAGATA 1260
 TGACAAAGTG GAAGAGATCA GCAGAAAGCA CAAACACAAC TGCTTCTGTA TTCAGGAGGT 1320
 65 TGTGAGTGGG CTGCGGCAGC CCGTGTGGTG CCTGCATAGT GGGGATGGCT CGCAACGTCT 1380
 CTTTCTCTG GAAAAGAGAG GTTATGTGAA GATACTTACC CTTGAAGGAG AATTTTTCAA 1440
 GAGGCTTTAT TTGACATTC ACAAACCTGT TCAAAGTGA ATAAAGGGAG GAGATGAAG 1500
 AGGACTGCTA AGCCTGCGAT TCCATCCCAA TTACAAGAAA AATGGAAGT TGTATGTGTC 1560
 70 CTATACCAAC AACCAAGAAC GGTGGGCTAT CGGGCCTCAT GACCACATTC TTAGGTTTGT 1620
 GGAATACACA GTATCCAGAA AAAATCCACA CCAAGTTGAT TTGAGAACAG CCAAGTCTT 1680
 TCTTGAAGTT GCAAGACTCC ACAGAAAGCA TCTGGGAGGA CAACCTGCTT TTGGCCCTGA 1740
 CGGCTTTTGT TACATCATTC TTGGTGTATG GATGATTACA CTGGATGATA TGGAAAGAA 1800
 GGAATGGTTA AGTGATTTC CAGGCTCAGT GCTACGGCTG GATGTGGACA CAGACATGTG 1860
 75 CAACGTGCTT TATTCATATC CAAGGAGCAA CCCACACTTC AACAGCACA ACCAGCCCCC 1920
 CGAAGTGTTC GCTCATGGGC TCCAGATCC AGGCAGATGT GCTGTGGATA GACATCCAC 1980
 TGTATATAAC ATCAATTTAA CGATACTGTG TTCAGACTCC AATGGAAGAA ACAGATCATC 2040

5 AGCCAGAAAT CTACAGATAA TAAAGGGGAA AGATTATGAA AGTGAGCCAT CACTTTTAGA 2100
 ATTCAAGCCA TTCAAGTAATG GTCTTTTGGT TGGTGGATTG GTATACCGGG GCTGCCAGTC 2160
 AGAAGAGATTG TATGGAAGCT ACGTGTITGG AGATCGTAAT GGGAAATTCC TAACCTCTCCA 2220
 GCAAAGTCTCT GTGACAAAGC AGTGGCAAGA AAAACCACTC TGCTCCGGCA CTAGTGGGTC 2280
 CTGTAGAGGC TACTTTTCCG GTCACATCTT GGGATTGGGA GAGATGAAC TAGGTGAAGT 2340
 TTACATTTTA TCAAGCAGTA AAGTATGAC CCAGACTCAC AATGGAAAAC TCTACAAAT 2400
 TGTAGATCCC AAAAGACCTT TAATGCCCTGA GGAATGCAGA GCCACGGTAC AACCTGCACA 2460
 GACACTGACT TCAGAGTGCT CCAGGCTCTG TCGAAACGCG TACTGCACCC CCACGGGAAA 2520
 GTGCTCTGCG AGTCCAGGCT GGGAGGGGGA CTCTGCAGA ACTGCAAAAT GTGAGCCAGC 2580
 10 ATGTCTCAT GGAGGTGCT GTGTAGACC GAACAAGTGC CTCTGTAATA AAGGATATCT 2640
 TGGTCTCAA TGTGAACAAG TGGACAGAAA CATCCGCGA GTGACCAAGG CAGGTATCT 2700
 TGATCAGATC ATTGACATGA CATCTTACTT GCTGGATCTA ACAAGTTACA TTGTATAGTT 2760
 TCTGGGACTG TTTGAATATT CTATTCCAT GGGCAITTTT TTTTATCTCT GTCAITAAA 2820
 AAAAAAGACT GTTATCTGTC TACACACTCC TGTGATTCA TTCTCTTTTA TTAATTTAAA 2880
 15 AATAATTTCC AGAAATGTGC AGATCCTCTG TGTGTATGTC AGCATGTTG TTCACATATG 2940
 CACATACACA TACTCATAAC CCTATATGTC GTTGTGTCAT AACAGATGAT TTTTAAAAAT 3000
 ATATACTTCC TTATGCAAG TAATTACAC AGAAATTCCA TTGTAATG ATAATGGATT 3060
 TTTTATGTTA CTAGAAGAGA TTATTGACT TCCACAGAAAT TTTCTGCTG TAATCACTAA 3120
 AGTCAACTTT AATAGAGTTT TGAACAGTA CTGTGCAATC CGATGGATCT AATTAAAAAA 3180
 20 AAGGCAATAT TTTTATATA AAGTACTATA CTAGGAGAGA ATGTTTCAGA ACTCCCTGAT 3240
 GAATTTCTAA GTGAGCAACT TGATATAAAA TTGTAATCTT CATTTTGTCT AGTGTATCCA 3300
 GTTACAGAAAT GCTACACACT TACCTTTTAA TTGGCTGAGA AATCTGGTTA TTTCACTTA 3360
 ATCTCAAGAT TGTTTTCAAG TGTTTTATAA TTAATCATA ATAGCATATT TAAAAATCAA 3420
 TCTTCTTAAA AGGTCTGCTT TTATTGTATA TTTTATTTAA CAATAGGCAG TGGGTTTGTG 3480
 25 TTACATATTT ATATATTTTA TTTTATTTT ATAATATAGA CATCACCTAG

Seq ID NO: 28 Protein sequence:
 Protein Accession #: XP_039209

30 1 11 21 31 41 51
 MLKMLSPKLL LLAVALGPF E GDAKFGERNE GSGARRRRL NGNPPKRLKR RDRRMSQLE 60
 LLSGGEMLCG GFYPRLSCL RSDSPGLRL ENKIFSVTN TECGLLEBI KCALCSPHSQ 120
 35 SLFHSPEREV LERDLVPLL CKDYCKEPFY TCRGHIPGL QTTADEPCFY YARDGGLCF 180
 PDFPRKQVRG PASNYLDQME EYDKVERISR KKHNCFCIQ EVVSGLRQPV GALHSGDGSQ 240
 RLFLILEKGY VKILTPGBI FKEPYLDIHK LVQSGIKGGD ERGLLSLAFH PNYKNGKLY 300
 VSVTTNQRRW AIGPHDHLR VVEYTVSRKN PHQVDLRTAR VFLEVAELHR KHLGGQLLFG 360
 40 PDGFLYILG DGMITLDDME EMDGLSDFTG SVLRDLVDTD MCNVPYSIPR SNPHFNSTNQ 420
 PPSVFAHGLH DPGRCVDRH FTDININLTI LCSDSNGKNR SSARILQIIK GKDYSESEPSL 480
 LEFKPFSNGP LVGGFVYRG QSERLYGSYV FGDRNGNFLT LQSPVTKQW QEKFLCLGTS 540
 GSCRGYFSGH ILGPGDELG EYILSSSKS MTQTHNGKLY KIVDPKRLPM PEBCRATVQP 600
 AQTLSSECSR LCRNGYCTPT GKCCSPGWB GDFCRATKCE PACRHGGVCV RPNKCLCKKG 660
 45 YLGPQCEQVD RNIRRVTRAG ILDQIIDMTS YLLDLTSYIV

Seq ID NO: 29 Nucleotide sequence:
 Nucleic Acid Accession #: NM_024756
 Coding sequence: 75..2924 (underlined sequences correspond to start and stop codons)

50 1 11 21 31 41 51
 AAGACAACGT CACTAGCAGT TTCTGGAGCT ACTTGCCAAG GCTGAGTGTG AGCTGAGCCT 60
 GCCCCACACC CAAGATGATC CTGAGCTTGC TGTTCAGCCT TGGGGGCCCC CTGGGCTGGG 120
 55 GGCTGTCTGGG GGCATGGGCC CAGGCTTCCA GTACTAGCCT CTCTGATCTG CAGAGCTCCA 180
 GGACACCTGG GGTCTGGAAG GCAGAGGCTG AGGACACCAG CAAGGACCCC GTTGGACGTA 240
 ACTGCTTGCCC CTACCCAATG TCCAAGCTGG TCACCTTACT AGCTCTTTGC AAAACAGAGA 300
 AATTCTCAT CATCTGCGAG CAGCCGTGTC CGCAGGGAGC TCCAGACTGC CAGAAAGTCA 360
 AAGTCAATGA CCGCATGGCC CACAAGCCAG TGTACCAGGT CAAGCAGAAG GTGCTGACCT 420
 60 CTTTGGCCTG GAGGTGCTGC CTTGGCTACA CGGGCCCAAA CTGCGAGCAC CACGATTCCA 480
 TGGCAATCCC TGAGCCTGCA GATCCTGGTG ACAGCCACCA GGAACCTCAG GATGGACCAG 540
 TCAGCTTCAA ACCTGGCCAC CTTGCTGCAG TGATCAATGA GOTTGAGGTG CAACAGGAAC 600
 AGCAGGAACA TCTGCTGGGA GATCTCCAGA ATGATGTGCA CCGGGTGGCA GACAGCCTGC 660
 CAGGCTGTG GAAAGCCCTG CTTGGTAACC TCACAGCTGC AGTGATGGAA GCAAAATCAA 720
 65 CAGGGCAGGA GTTCCCTGAT AGATCTTTGG AGCAGGTGCT GCTACCCACG GTGACACCT 780
 TCTTCAAGT GCATTTGAG CCAATCTGGA GAGAGCTTAA CCAAGCCTG CACAGCTTGA 840
 CCCAGGCCAT AAGAAACCTG TCTCTTGACG TGGAGGCCAA CCGCCAGGCC ATCTCCAGAG 900
 TCCAGGACAG TGCCGTGGCC AGGGCTGACT TCCAGGAGCT TGGTGCCAAA TTTGAGGCCA 960
 AGGTCACAGA GAACACTCAG AGAGTGGGTC AGCTGCGACA GGACGTGGAG GACCGCCTGC 1020
 70 ACGCCACGCA CTTTACCCTG CACCGCTCGA TCTCAGAGCT CCAAGCCGAT GTGACACCA 1080
 AATTGAAGAG GCTGCACAAG GCTCAGGAGG CCCCAGGAC CAATGGCAGT CTGCTGTTGG 1140
 CAACGCTCTG GGCTGGGCA AGGCTGTAGC CGGACAGCCT GCAGGCCAGG CTGGGCCAGC 1200
 TGCAGAGGAA CTTCTCAGAG CTGCACATGA CCAAGGCCCG CAGGAGGAG GAGTTGAGT 1260
 75 ACACCCCTGA GGACATGAGG GCCACCTGGA CCGGCAAGT GGATGAGATC AAGGAACTGT 1320
 ACTCCGAATC GGACGAGACT TTCGATCAGA TTAGCAAGGT GGAGCGCGAG GTGAGGAGC 1380
 TGCAGGTGAA CCACACGGCG CTCGCTGAGC TGGCGGTGAT CCTGATGGAG AAGTCTCTGA 1440

5 TCATGGAGGA GAACAAGGAG GAGGTGGAGC GGCAGCTCCT GGAGCTCAAC CTCACGCTGC 1500
 AGCACCTGCA GGGTGGCCAT GCCGACCTCA TCAAGTAAGT GAAGGACTGC AATTGCCAGA 1560
 AGCTCTATTT AGACCTGGAC GTCATCCGGG AGGGCCAGAG GGAAGCCAGG GGTGCCCTGG 1620
 AGGAGAGCCA GGTGAGCCTG GACGAGCGGC GGCAGCTGGA CGCTCCTCC CTCGAGGCC 1680
 10 TGCAAGACCG CTGTGACGCC GTGTGCTGG CCGTGGACGC GCACAAAGCG GAGGGCGAGC 1740
 GGGCGCGGGC GGGCAGCTCG CGGCTCCGGA GCCAAGTGCA GCGCTGGAT GACGAGGTGG 1800
 GCGCGCTGAA GCGCGCGCGG GCGGAGGCC GGCACGAGGT GCGCCAGCTG CACAGCGCCT 1860
 TCGCGCGCCT GCTGGAGGAC GCGCTGCGGC ACGAGGCGGT GCTGGCGCGG CTCTTCGGGG 1920
 AGGAGGTGCT GGAGGAGATG TCTGAGCAGA CGCCGGGACC GCTGCCCTGG AGCTACGAGC 1980
 15 AGATCCCGCT GGGCTGCGAG GACGCGCTA GCGGGCTGCA GGAGCAGGCG CTCGGCTGGG 2040
 ACGAGCTGGC CGCCGAGTGT ACGGCGCTGG AGCAGGCGCT GGAGCCCCG GCGCGCGCAG 2100
 AGCACCTGGA GCGCAGCCAC GACGCGGGCC GCGAGGAGGC CGCCACCACC GCGCTGGCGG 2160
 GSGTGGCGCG GAGCTCCAG AGCCTGAGCA ACGACGTCAA GAATGTCGGG CGGTGCTCGG 2220
 AGGCGGAGCG CGGGCGCGGG GCGGCGCTCC TCAACGCTCC CTTGAGCGG CTCACCAACG 2280
 20 GGAATCTTGC CACTCAGCGC AGCTTGGAGC AGCACCAGCG GCTCTTCCAC AGCCTCTTTG 2340
 GGAATCTTCA AGGCTCTATG GAAGCCAAAG TCAGCTTGGA CTTGGGGAAG CTCGAGACCA 2400
 TGCTGAGCAG GAAAGGGAAG AAGCAGCAGA AAGACCTGGA AGCTCCCCG AAGAGGGACA 2460
 AGAAGGAAGC GGAACCTTTG GTGAGCATAC GGGTCACAGG GCGTGTGCCA GGTGCTTGG 2520
 GCGCGCGCCT CTGGGAGGCA GGATCCCTTG TGGCTTTCTA TGCCAGCTTT TCAGAAGGGA 2580
 25 CGGCTGCGCT GCAGACAGTG AAGTTCAACA CCACATACAT CAACATTGGC AGCAGCTACT 2640
 TCCCTGAACA TGCTACTTTC CGAGCCCTTG AGCGTGTGT CTACCTGTTT CGAGTGAGGG 2700
 TTGAATTTGG CCCAGGCCCA GGCACCGGGC AGCTGGTGTG TGGAGTGCAC CATCGGACTC 2760
 CAGTCTGTAC CACTGGGCGG GGGAGTGGAA GCACAGCAAC GGTCTTTGCC ATGGCTGAGC 2820
 30 TGCAAGAGGG TGAGCGAGTA TGGTTTGAAT TAACCCAGGG ATCAATAACA AAGAGAAGCC 2880
 TGTGCGGCAC TGCAATTTGG GCGTTCCTGA TGTTTAAGAC CTGAACCCCA GCCCAATCT 2940
 GATCAGACAT CATGGACTCG CCCAGCTCTC CTGCGCCTGG GCGTCTGGCC AAGGATGGGC 3000
 TGGAGGTGAT TCAGTTGGTC TGTCTCTTCC CTGGAACCTC TCTGCAAGA TGGTGTGGTG 3060
 TAGCTGGCTT CCCTGTAAAC ACATGGGGCT TGGCCATTTC TCCATGATGA GAAGGACTGG 3120
 35 AATGCTCTTC CGGCGAGGAC ATGGTCTTAG GAAGCTGAA CTTGGCTTG GCATGCCTTC 3180
 TCAGACAGCA CGGCTGGGC TCCAACCTCT CACCAACACC TGTATTCTAC AACTTCTTTG 3240
 GTGTTTGTCT CCTCTGTGG TTGGAACCTT CTGTACAACA CTTTAACTT TTCTCTTGCT 3300
 TCCTCTCTTC TTCTCCCTTA TCGTATGATA GAAAGACATT CTTCGCCAGG AGGAATGTTT 3360
 40 AATATGGAGG CAACATTTTG GCCAACATTG GAAAGCACTA GAGGGCAATG GGATTTAAACC 3420
 AACCTGCTTG GTCTCTATTA GTCAATTAAT AAGACGACAG CCGTGCCAC CAGGGGAAG 3480
 GAATATAGTA TCTTTAGTTT CAGTCAATCC TTGTAGGATA TGGTTTAGCT GTGCCCCAC 3540
 CTAAATATAT ATCTTGAATT GTAATCCCTA TAATCCCCAC ATCAAGGGAG AGATCAGGTG 3600
 TAGGTAAATT GATCTTGGGG GCGTTTCCCC CATGCTGTTT TTGTGATAGT TCTCAGGAGA 3660
 45 TCTGTATGAT TTATAAGTTT GATAGTTCTT CCTGTGTTCA TTCTCTTCC TGCCACCTTG 3720
 TGAAGATGCC TTGTTCTCTC TCACTGTCT GCCATGATTG TAAGTTTCTT GAGGCTTCCC 3780
 CAGCCATGTG GAACAGTGAG TCAATTAAAC CTCTTCTCTT TATAAAT

Seq ID NO: 30 Protein sequence:
 Protein Accession #: NP_079032

45 1 11 21 31 41 51
 MILSLPLSLG GPLWGLLGA WAQASSTSLD DLQSSRTPGV WKAEDTSTK DPGVGRNWCYPY 60
 50 PMSKLVTLIA LCKTEKFLIH SQQPCPQCAP DQKVKVMYR MAHKPVYQVK QKVLTSIAWR 120
 CCPGYTGPMC EHHDSMAIPE PADPGDSHQE PQDGPVSEFKP GHAAVINEV EVQGEQGEHL 180
 LGDLQNDVHR VADSLPGLWK ALPGNLTAAV MEANQTGHEF PDRSLEQVLL PHVDTFQVH 240
 FSPINRSFNQ SLHSLTQAIR NLSLDVEANR QAIRSVQDSA VARADPQELG AKFEAKVQEN 300
 55 TQRVGQLRQD VEDRLHAQHF TLHRSISBLQ ADVDTKLKRL HKAQERAPGTN GSVLVIATPGA 360
 GARPEPDSLQ ARLQLQRLNL SELHMTTARR BEELQVLTED MRATLTRHVD EIKELYSHSD 420
 ETPDQISKVE RQVSELQVNH TALRELRVIL MEKSLIMEEN KBEVERQLLE LNLTLQHLQG 480
 GHADLIKVKV DCNCQKLYLD LDVIREGQRD ATRALESTQV SLDERRQLDG SSLQALQNAV 540
 60 DAVSLAVDAH KAEGERARAA TSRLRSQVQA LDDEVGALKA AAREARHEVR QLHSAFAPALL 600
 EDALRHEAVL AALPQEVLE EMSEQTGPPL PLSYEQIRVA LQDAASGLQE QALGWDELA 660
 RVTALEQASE PPRPAHLEP SHDAGREBAA TTALAGLARE LQSLSNVKN VGRCCBAEAG 720
 AGAASLNASL DGLHNLATP QRSLEQHRL PHSLEFNGFG LMEANVSLDL GKLQTMISRK 780
 GKQKQKLEA PRKRDKKBAE PLVDIRVTGP VPGALGAALW EAGSFVAFYA SPSEGTAAIQ 840
 TVKFNNTYIN IGSSYFPEHG YFRAPERGVY LPAVSVEFGP GPQTQLVFG GEHRTFVCTT 900
 QGSGSSTATV FAMAELQKGB RVWFELTQGS ITRSLSGTA FGGFLMFKT

65 Seq ID NO: 31 Nucleotide sequence:
 Nucleic Acid Accession #: AB037715
 Coding sequence: 370..3489 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 GAACGCTCAC AGAACAGGCA GTGCAATTCC ATGTTCTCTT TAAATATGTT AGCCCTACCG 60
 GAGCTGAGC TGGCCAGTCT ACTTGGAGAG GAAAGTAGA TCTGGGGAAG GTGGAAGGGT 120
 75 CAGTTCTTAA GTGACTTCTT CCTCGGGGAT GGTAAGGGCA TTTGCTGATC TCCAGTGACT 180
 CCTCGGTGCC TCATGTCAG ACTCGGCTGT CTCACCTCCA GATATCTGAT TTTGCAAAA 240
 GGGACACACC TATCTGCAGC AAAGAAGACA CTGACCAGAT TCGCAGCGGT GCTTTTGGAT 300

	GCTCTGTAGC	CACCCGGGGC	CCAGGAGGAC	TGACTCGGCA	GCAGGATTGG	TGCAATGGGA	360
	TCGGAGACCA	TGGCAGTGCA	GCTGGTGGCC	GACTCAGCTC	TCGGCCTGCT	GATGATGACG	420
	GAGGGCCGCC	GATGTCAAGT	ACATCTTCTT	GATGACAGGA	AGCTGGAACT	CCTAGTACAG	480
	CCCAAGCTGT	TGGCCAAAGG	GCTTCTTGAC	CTTGTGGCTT	CTCACTTCAA	TCTGAAGGAA	540
5	AAGGAGTACT	TTGGAAATAGC	ATTACAGAT	GAACCGGAC	ACTTAACTG	GCTTCAGCTA	600
	GATCGAAGAG	TATTTGGAACA	TGACTTCCCT	AAAAAGTCAG	GACCCGTGGT	TTTATACITT	660
	TGTCAGAGGT	TCTATATAGA	AAGCATTTCA	TACCTGAAGG	ATAATGCTAC	CATTGAGCTT	720
	TTCTTTCTGA	AOCGGAAGTC	CTGCATCTAC	AAGGAGCTTA	TTGAOGTTGA	CAGOGAAGTG	780
10	GTGTTTGAAT	TAGCTTCTTA	TATTTTACAG	GAGGCAAAAG	GAGATTTTTC	TAGCAATGAA	840
	GTTCGTAGGA	GTGACTTGAA	GAAGCTGCCA	GCCCTTCCCA	CCCAAGCCCT	GAAGGAGCAC	900
	CCTTCCCTGG	CCTACTGTGA	AGACAGAGTC	ATTGAGCACT	ACAAGAACT	GAACGTCAG	960
	ACAAGAGGTC	AAGCAATGCT	AAACTACATG	AGCATCGTGG	AGTCTCTCCC	AACCTACGGG	1020
	GTTCACATTT	ATGCAATGAA	GGACAAGCAG	GGCATACCAT	GGTGCTGGGG	CCTGAGCTAC	1080
	AAAGGATCT	TCCAGTATGA	GTACCATGAT	AAAGTGAAGC	CAAGAAAGAT	ATTCCAAATG	1140
15	AGACAGTTGG	AAAACTTTTA	CTTCAGAGAA	AAGAAGTTTT	CCGTGGAAAT	TCATGACCCA	1200
	CGCAGGGCTT	CAGTGACAAG	GAGGACGTTT	GGGCACAGCG	GCATTGCACT	GCACACGTGG	1260
	TATCATGTTC	CGGCATTGAT	CAAGTCCATC	TGGGCTATGG	CCATAAGCCA	ACACCAATTG	1320
	TATCTGGACA	GAAGACAGAG	TAAGTCCAAA	ATCCATGTCG	CACGCAGCCT	GAGTGAATTC	1380
20	GCCATCGACC	TGACCGAGAC	GGGACGCTGG	AAGACCTCGA	AGCTGGCCAA	CATGGGTAGC	1440
	AAGGGGAAGA	TATCAGCGG	CAGCAGCGGC	AGCCTGCTGT	CTTCAGGTTC	TCAGGAATCA	1500
	GATAGCTCCG	AGTGGGCCAA	GAAGGACATG	CTGGCTGCTT	TGAAGTCCAG	CAGGAAAGCT	1560
	CTGGAGGAAA	CCCTCGTCA	GAGGCTGGAG	GAATGAAGA	AGCTGTGTCT	CCGAGAAGCT	1620
	GAGCTCACGG	GCAAGCTGCC	AGTAGAATAT	CCCTGGATC	CAGGGGAGGA	ACCAACCATT	1680
25	GTTCGGAGAA	GAATAGGAAC	AGCCTTCAAA	CTGGATGAAC	AGAAAATCCT	GCCCAAAGGA	1740
	GAGGAAGCTG	AGCTGGAAAG	CCTGGAAAGA	GAGTTTGCCA	TTCACTCCCA	GATTACGGAG	1800
	GCCGCCCGCG	GCTTAGCCAG	TGACCCCAAC	GTACGCAAAA	AACCTGAAGA	ACAAAGGAAA	1860
	ACCTCGTATC	TGAATGCACT	GAAGAAACTG	CAGGAGATTG	AAAAATGCAAT	CAATGAGAAC	1920
	CGCATCTCCG	CTGGGAAGAA	ACCCACCCAG	AGGGCTTCGC	TGATCATAGA	CGATGGAAAC	1980
	ATTGCCAGTG	AAGACAGCTC	CCTCTCAGAT	GCCCTTGTTC	TTGAGGATGA	AGACTCTCAG	2040
30	GTATACAGCA	CAATATCCCC	CCTACATTTCT	CCTCAAGAG	GACTCCCTCC	TCGGCCACCG	2100
	TCGCACAACA	GGCTCTCTCC	TCCCACTGCC	CTGGAGGAGC	TCCACAGAT	GCATATCAC	2160
	CGCAACGACT	ATGACAAGTC	ACCATCAAG	CCCAAAATGT	GGAGTGAGTC	CTCTTTAGAT	2220
	GAACCTCTAT	AGAAGTCAA	GAAGCGCTCC	TCTCACAGCC	ATTCAGCAG	CCACAGCGC	2280
35	TTCCCCAGCA	CAGGAAGCTG	TGCGGAAGCC	GGCGGAGGAA	GCAACTCCTT	GCAGAACAGC	2340
	CCCATCCGCG	GCTCCCGCA	CTGGAATCCC	CAGTCCAGCA	TGCGCTCCAC	GCCAGACCTG	2400
	CGGTCGGGA	GTCCCACTA	CTGCTCAATC	ACGAGGTCGG	TGGACATCAG	CCCCACCCGA	2460
	TCGCTGAGCA	TGCACTGCA	CTTTAGGCAC	CGGAGCTCCA	GCTTGGAGTC	CCAGGGCAAG	2520
	CTCCTGGGCT	CGGAAACGA	CACCGGAGC	CCGACTTCTC	ACACCCCGCG	GACTCGTAGC	2580
	AGCAACGGCT	CAGACCCCAT	GGACGACTGC	TCGTCTGTGA	CCAGCCACTC	GAGCTCGGAG	2640
40	CACCTACTAC	CGGCGCAGAT	GAACGCCAAC	TACTCCACGC	TGGCCGAGGA	CTCGCCGTCC	2700
	AAGGCGCGCC	AGAGGCGAG	GCAGCGGCG	CGGCGGCGG	GCGCACTGGG	CTCAGCCAGC	2760
	TCGGGCGAGC	TGCCCAACT	GGCGGCGCG	GGGGGTGCG	GGGGCGCGG	GGGCGCGGG	2820
	CGCGGTGTAT	ACCTGCACAG	CCAGAGCCAG	CCAGCTCGC	AGTACCGCAT	CAAGGAGTAC	2880
45	CCGCTCTGCA	TCGAGGCGG	CGCCACGCC	GTGGTGGTGC	GCAGCTTGA	GAGCGACGAG	2940
	GAGTGCCACT	ACAGCGTCAA	GGCTCAGTTC	AAGACGTCCA	ACTCTACAC	GGCGGGCGGC	3000
	CTGTTTCAAG	AGAGCTGGCG	CGGCGGCGG	GGCGACGAGG	GCACACGCG	CGCGCTGACG	3060
	CCGTGCGGAT	CGCAGATCTC	CGGAGCTCG	TCGTGGGCC	CGGAGGGCGC	CCAGCAGAAC	3120
	GGCGCGGGCC	GTGCGCGGT	CTCAGACGAG	CTGCGCCAGT	GATACAGCG	TTCCACCGCC	3180
50	TCGCACAAGG	AGCACAAGCG	CCTGTGCGAC	ACCAGCTCCA	CCTCTCGGA	CAGGGGCTCG	3240
	CAGTACAGCA	CCTCTCCCA	GAGCACCTTC	GTGGCGCACA	GCAGGGTCAC	CAGGATGCC	3300
	CAGATGTGCA	AGGCCAAGTC	AGCTGCCTTA	CCTCAAAGCC	AGAGAAGCTC	GACACCGTCA	3360
	AGTGAAATTG	GAGCCACCCC	CCCAAGCAGC	CCCCACACCA	TCTTAACCTG	GCAGACTGGA	3420
	GAAGCAACAG	AAAACTCACC	CATTCTGGAT	GGGTCTGAGT	CTCCACCTCA	CCAAAGTACT	3480
55	GATGAATAGA	GGAGCTACAA	TGATAGCTGT	TTCTTGGATT	CCTCCCTCTA	TCCAGAACTA	3540
	GCTGATGTCC	AGTGTACCG	GCAGGAAAAA	GCCAAAGCCG	GGACCTCGT	GTGAGCCAGC	3600
	CCGCGCTAAT	CTGACCGCCT	CAACGCCAAT	CTGAGATCAC	CTCACTGCCT	CTCAATTGCC	3660
	TTACCCAGAT	GCAACGTCAC	CCTGCACCAG	CTTTGGCCCT	CAGCACTTTT	TTTCTCTGTT	3720
	CTCCGCAATC	CCTCCCTCTT	GAACCACTGA	CTGAGGAGAC	ATTCTGGAGG	GTTCGGGTCC	3780
60	CACCTGTGTG	CCCTTGGCG	TCTTGGCCAT	AGAGAGCCAG	ACACCAATCC	TCAATGGCAC	3840
	CTTGTGTGGT	TCCCTTGCC	ATGACAGCCC	CTAGGCCAGG	AACCATCAGG	GGGCGCAGCC	3900
	GGCATCCAAT	TCTCGCGGAT	AAGTAGCGTT	GGGAGAGAAC	GGGAAAGGGG	ACTTGGGTTA	3960
	CAGGGTGACC	CAGAAAGAGC	ATTGAGCTGT	GTCCAGCCTG	CCACCATATC	GTAGGCCAAC	4020
	CAAGCACTTC	ATGAAGAGGA	GGCCTCGTGG	CATATTTCAGT	TTACACCTGA	AATATTCTCT	4080
65	GATGGGACAG	CTTGTGGGA	TGGCTATGGG	GGAAGGGGAG	GTGAGAAAG	GAGTTCTCG	4140
	TAAACAGAAA	TGATCGGAG	GACCAACATC	AGTTCTATGC	TGCCAAAGAT	TAAATAATAA	4200
	TAACCAATTA	AAAAATTAAG	AGGGGCCAAG	AGGAAGACAT	TCTTTCTGCA	AGGAATTTTC	4260
	TTTAAATTC	TGAATGCTA	CTACACACNA	GTAAAGTCA	ACCCTATGTA	AATGTTGTC	4320
	CTCTCTCTAG	CCCTCTCCCT	TACTGGCCCA	CTTCTCTCTC	CGTAGAGAGC	CTGAAAACT	4380
70	GCCCAATGC	CACGTAAGG	GCGAGGAAGT	CTTGGCTGGC	GTGCTGACT	CACAGTCGCC	4440
	ATCCATCTGG	ACACAAGAG	AGACCTGTGG	GAGTCATAGA	GGGTACTGTT	AGCCCGGTC	4500
	CTCAGAGGCT	GTTCAGCGA	GCCCAAGACT	CAAGCTGCT	TTCTTTTCAG	GATTTGTAGT	4560
	AACGTAAAGT	GATAATGGCC	AAAAGTGGTT	CTCTCTCATT	AAACCAACCA	GTAAGAGCGT	4620
	ATCCTATTTT	TTTGCAATAG	GTGTTTTCATT	TTGGTTTTTA	TGGGAAACCA	AGGGAAAGCG	4680
	ACATTGGCAT	CCATTGAGT	TTTAACTGTC	GTGGCTCATT	TTCTGTTCTG	TAGCACTTGT	4740
75	GTGACAAAAG	AGCTCAGATC	CGACTTCTCC	TATGTGTCAC	TTATTCCAAG	AACCCAACCTA	4800
	TGCCCTTAGG	TAGAAAGATT	TGACTCGTGT	GTCTACTAGC	CAACAGGCAG	AGCAGGTTTG	4860

5
 10
 15
 20
 25
 30
 35

```

AAAAAATAT CAGCTCCCAA AGGGCCCATG TGTCTACATC ATCAGTTACT GTCATGCACC 4920
ACATTTGTGT GCAGATACCA AAAGAGGAGG AAAGAAGAAA AAAATTAATG TGTGGGAGCT 4980
GCACGTTTAC ATGTTTTGAG CTATGCTTCA AACACAACTG GAAAGCCATC AATCTTCAAA 5040
GGCCTCAAAA ATACTTTTAT AGTAACAAGT GCACGACTTT AGTTGGGTTA TTCAAGATGG 5100
CACAAAAAGG TTTCGGCAGA GGTGGTATGC TGTGCTTTTG GCGCAAGTGG TGGGGGGATG 5160
GGGGTGGGGG TGGAAATTTT TTCTCACTCT AATGACTTCC TATTGGAAGG GCATTGACAG 5220
CCAGGGACAG GAGCCAGGGT GGGGGTAGTT TTGTGGGAAA GCAGAACTGA AGTTAGCTTA 5280
AGCATAAAAA CAAAGAAAAA TCTTCGCTTT TCATGTATGT GGAATCCAAG AATAACCAT 5340
GGCTCTACCA GACCAGGAGG GTAAGGATGG ACACATAAAT GAAACAAATA CCAAGGTATT 5400
CCTTCTGCTG CAGCCTGGAG ACCACCGAGA GTGAGCTGGG GGCACACACA CACCTGGCGG 5460
GGACCCGGCA GGGACAAGGC GGGCCGTGGC CTCTCCACCC AAGTCTCTCT AGACAATTCA 5520
GGCCCTGCTT TCCCAGCTCT CATGCATGGC TGGACTGGTG ATTCCAGGGT GCAGAAGGGA 5580
TTCATATTCC CAGAACGCTT TAAGTGTACA CCTGCAGGAT AAAGAGATAC CGGTTACATT 5640
ATTAAATAGT TCTAGGATTT CACTGGGGGA TATTTTGTGT GCTTTTACTT TCATGGTTAG 5700
AGCTACAAAG AACAGTGATT TTTTTTTTTT CTCCCTTCCC CATTGAGAAA CATTATACAT 5760
TGGGCCATTT TTCTTCTCC CAAAGAAGAT TCATGGATAG TCAGACTGAA CTGTGTGCAA 5820
CAGGAAAAAG CAAAGGGGAA AAGGCAGCTG ATGAGGTTAC ATGGTTACAT GTTCTACATC 5880
ATCAGAGATG GCTTGAATAT TAGTCTGGAG AAAATGGGAT CAGGATTCTA GCCCACTGGA 5940
GTTCAGAGGA ATGAGAGGCA AAAATTCTAA AGATTTGGGT TATATTTTCA ACTTGGGGGA 6000
CAGAGAGAAA TGGAGAGCAG GAATTACAGT TCACAACAAAC ATCATGATAG TCTGTGTAGT 6060
AAGACAGAGA TTAAGTAAAA CAGGTTTTAC TGTTTAGCTG AGTTCAGTTA ATACAAAATG 6120
TACATAAAAC GTTAGTCTCT TGAGACTGAC ATGATTAAATG ATCAGTGTGG TGGGAAATGA 6180
TGTAGTTAAT GTACACAAGC ACTTGCACAA TCTTTATCCC TATTTCTTTA AAACAAAATA 6240
AGGTGAAATA CGAAGTCTCT GGTCTGATAT AAAGCCCTTA TTGATTCTCT CGGATGCGTA 6300
AAGAAAAATG CCTGTTTTCAG CCAGAAAGACT GGTGAAAACA CATACATCAG ACTATGTTGT 6360
GAGCCAGGAT GATTTTATAT TTTATATAT GCAGGTGAGT GTTGAAACTG TTAATAATCC 6420
AATTTGTTT CATTCAGTAT TAGTTTAGTT CTAATATAG CAAACCCCAT CCAGGTGCTA 6480
TCAGATGACC AGTTACTGCT TAGTTAATA GGTGTAAGT TTTACATATA CATTAAATTC 6540
AATAGTTTAT TACAAGTTGT GTAAATGGA CTCTAGTTTA ATAATGGGGG AAAAAGATT 6600
AGGTGTGCTC TGAACCTGAC TGTAGAGCAT GTAAATGAT TTTACTGGAT TCTGTCTCAC 6660
TGTAAATCAAT GAAAAAGATG TACGTTGTAG ACAAGATTGC AGAATTAAAA AAAGAAATCT 6720
GCTTTTAATT TATCTTTTTT GTATTAGAA TTTGTATAGT ATCTTTACAT TTTGCAAAAC 6780
AGTGTGTGTA CACACTTATTA AAGCATTTTC AAAATG
  
```

Seq ID NO: 32 Protein sequence:
 Protein Accession #: BAA92532

40
 45
 50
 55
 60

```

1 11 21 31 41 51
MAVQLVPDPA LGLLMTTEGR RCQVHLDDR KLELLVQPKL LAKELLDLVA SHPNLKEKEY 60
FGIAPDTBTG HLNWLQDRLR VLEHDFPKKS GPVVLYFCVR FYIESISYLY DNATIELPFL 120
NAKSCIYKEL IDVDSEVVF E LASYILQBAK GDFSSNEVVR SDLKLPALP TQALKEHPSL 180
AYCEDRVIEH YKKLNGQTRG QAIVNYSIV ELSLPTGVHY YAVKDKQGP WMLGLSYKGI 240
FOYDVKDKVK PRKIPQWRQL ENLYPREKKP SVEVHDPERR SVTRTTPGHS GLAVHTWYAC 300
PALIKSIWAM AISOHQFYLD RKQSKSKINA ARSLSLAID LTETGTLKTS KLANMGSKGK 360
IISGSSGSL SSGSQSDSS QSAKKDLAA LKSRQBALEE TLRQRLBZLK KLCRLBZLT 420
KGLPVEYFLD PGEEPIVRR RIGTAPKLDE QKILPKGEA ELERLEREFA IQSQITEAAR 480
RLASDPNVSK KLKQKRKTSY LNALKKLQEI ENAINENRIK SGKKPTQRA S LIIDGNIAS 540
EDSSLSDALV LEDEDSQVTS TISPLHSPHK GLPPRPSSH N RPPPPQSLBG LRQMHYHRND 600
YDKSPIKPKM WSESSLDEPY EKVKRSSH S HSSSHKRPFS TGSCAEAGG S SLSLQNSPIR 660
OLPHWNSQSS MPSTPDLVR SPHYVHSTRS VDISPTRLHS LALHFRHRS S SLESQKLLG 720
SEMDTGS PDP YTPTRSSNG SDPMDDCSS TSHSSSEHY Y PAQMANYST LAEDSPSKAR 780
QRQRQRQRAA GALGSASSG MPNLAARGGA GGAGGAGG V YLHSSQSPSS QYRIKEYPLY 840
IBGGATPVVV RSLESDBECH YSVKAQPKTS NSYTAGGLEK ESWRSGGGDE GDTGRLTPSR 900
SQUILTPSLG REGAHDKAG RAAVSDSLRQ WYQRTASHK EHSRLSHTSS TSSDSGSQYS 960
TSSQSTFVAH SRVTRMPQMC KATSAALPQS QRSSTPSSEI GATPFSSPHR ILTWQTGEAT 1020
ENSPILDGSE SPPHQSTDE
  
```

Seq ID NO: 33 Nucleotide sequence:
 Nucleic Acid Accession #: NM_014331
 Coding sequence: 1..1506 (underlined sequences correspond to start and stop codons)

65
 70
 75

```

1 11 21 31 41 51
ATGTCAGAA AGCCTGTGT GTCCACCATC TCCAAAGGAG GTTACCTGCA GGGAAATGTT 60
AACGGGAGGC TGCCTTCCCT GGGCAACAAG GAGCCACCTG GGCAGGAGAA AGTGCAGCTG 120
AAGAGGAAAG TCACCTTACT GAGGGGAGTC TCACATTATCA TTGGCACCAT CATTGAGACA 180
GGAATCTTCA TCTCTCTAA GGGCGTGTCT CAGAACAAGG GCAGCTGGG CATGTCTCTG 240
ACCATCTGGA CGGTGTGTGG GGTCTGTGTA CTATTGGAG CTTTGTCTTA TGCTGAATTG 300
GGAACAACTA TAAAGAAATC TGGAGGTGAT TACACATATA TTTTGAAGT CTTTGTGCTA 360
TTACCACTCT TTGTACGAT CTGGGTGGAA CTCTCATATA TACGCCCTGC AGCTACTGCT 420
GTGATATCCC TGGCATTGG ACCTACATT CTGGAACCAT TTTTATTATCA ATGTGAATC 480
  
```


5 CCGAACTTG OGATCAAGCT CATTACAGCT GTGGGCATAA CTGTAGTGAT GGTCTAAAT 540
 AGCATGAGTG TCAGCTGGAG CGCCCGGATC CAGATTTTCT TAACCTTTTG CAAGCTCACA 600
 GCAATTCTGA TAAATTATAGT CCTGGGAGTT ATGCAGCTAA TTAAGGTCA AAGCAGAAC 660
 TTTAAGACG CGTTTTCAGG AAGAGATTCA AGTATTACGC GGTGCCACT GGCCTTTTAT 720
 TATGGAATGT ATGCATATGC TGGCTGGTTT TACCTCAACT TTGTACTGA AGAAGTAGAA 780
 AACCTGAGAA AAACCATTCG CCTTGCATAA TGTATATCCA TGSCCATGT CACCATTTGC 840
 TATGTGCTGA CAAATGTGGC CTACTTTACG ACCATTAAAT CTGAGGAGCT GCTGCTTTCA 900
 AATCAGTGG CAGTGAOCTT TTCTGAGCGG CTACTGGGAA ATTTCTCATT AGCAGTTCCG 960
 10 ATCTTTGTTG CCTCTCCTG CTTTGGCTCC ATGAAOGGTG GTGTGTTTGC TGTCTCCAGG 1020
 TTATCTATG TTGCGTCTCG AGAGGGTCAC CTTCAGAAA TCCTCTCCAT GATTCAATGTC 1080
 CGCAAGCAGC CTCTCTACC AGCTGTTATT GTTTTGACCC CTTTGACAA GATAATGCTC 1140
 TTCTCTGAG AGCTCGACAG TCTTTTGAAT TTCTCTCAGT TTGCCAGGTG GCTTTTATT 1200
 GGGCTGGCAG TTGCTGGGCT GATTATCTT CGATACAAAT GCCAGATAT GCATCTCTCT 1260
 TTCAAGGTGC CACTGTTTAT CCCAGCTTTG TTTTCTCTCA CATGCTCTT CATGGTTGCC 1320
 15 TTTTCTCTCT ATTCGGACCC ATTTAGTACA GGGATTGGCT TCGTCATCAC TCTGACTGGA 1380
 GTCCGTGGCT ATTTCTCTT TATTATATGG GACAGAAAC CCAGTGTTT TAGAATAATG 1440
 TCAAGAGAAA TAACCAAGAC ATTACAAATA ATACTGGAAG TTGTACCAGA AGAAGATAAG 1500
 TTATGACTA ATGGAATGGA GATCTTGGCA ATCTGCCCAA GGGAGACAC AAAATAGGGA 1560
 TTTTACTCTC ATTTCTGAA AGTCTAGAGA ATTACAACTT TGTGATAAA CAAAAGGAGT 1620
 20 CAGTATTATT TATTATATA TTTTAGCATA TTGAACTAA TTCTAAGAA ATTTAGTTAT 1680
 AACCTATGT AGTTATAGAA AGTGANTATG CAGTTATTCT ATGAGTCGCA CAATCTTGA 1740
 GTCTCTGATA CCTACCTATT GGGGTAGGA GAAAAGACTA GACAATTACT ATGTGTCAT 1800
 TCTCTACAC ATATGTTAGC ACGGCAAGA ACCTTCAAA TGAAGACTGA GATTTTCTG 1860
 25 TATATATGGG TTTGTAAAG ATGGTTTAC ACACATACAGA TGTCTACT GTGAAAAGTG 1920
 TTTTCAATCT TGAAGAAAG CATACATCAT GATTATGGCA AAGAGGAGAG AAGAAAATT 1980
 ATTTTACATT GACATTCAT TGCTTCCCT TAGATACCAA TTAGATAAC AAACACTCAT 2040
 GCTTTAATGG ATTATACCCA GAGCACTTTG AACAAAGTCT AGTGGGATG GTTGAATACA 2100
 TTAAAGAGA GTTCTAGGG GCTACTGTTT ATGAGACACA TCCAGGAGT ATGTTAAGT 2160
 AAAAATCTCT GAGAATTAT TATGTCAGAT GTTTTTCAT TCATTATCAG GAAGTTTAG 2220
 30 TATCTGTCA TTTTCTTTT TCACATCAGT TTGATCAGGA AAGGTATAA CACATCTTAG 2280
 AGCAAGAGTT AGTTTGTAT TAAATCTCA TTAGAACAAC CACTGTTTC ACTAATAACT 2340
 TACCCCTGAT GAGTCTATCT AAACATATGC ATTTTAAAGC TTCAAATAC ATTATCAACA 2400
 TGAGAGAAAT AACCAACAAA GAAGATGTTT AAAATAATAG TCCATATCT GTAATCATAT 2460
 35 CTACATGCAA TGTATGTAAT TCTGAAGTTT TTTAAATTA TGGCTATTT TACACGATGA 2520
 TGAATTTTGA CAGTTTGTGC ATTTTCTTTA TACATTTTAT ATTCTCTGT TAAATATCT 2580
 CTTCAATGTA AACTGTCCAG ATTAATTAGG AAAAGGCATA TATTACATA AAAATTGCAA 2640
 AAGAAATGTC GCTGTAAATA AGATTTACAA CTGATGTTTC TAGAAAAATT CCACCTCTAT 2700
 ATCTAGGCTT TGTAGTAAT TTCCACACCT TAATTATCAT TCACTTGCA AAGAGACAA 2760
 40 GTGTAAGAA GAAAATGAA ATGAGAATCT GTGGATAAGT GTTGTGTTT AGAAGATGTT 2820
 CTTTTGCGAG TATTAGAAAA TACTGTGAGC CGGGCATGGT GGCTTACATC TGTAAATCCA 2880
 GCACCTTGGG AGGCTGAGGG GGTGATCAC CTGAGGTGGG GAGTTCTAGA CCAGCTGAC 2940
 CAACATGGAG AAACCCATC TCTACTAAAA ATACAAAATT AGCTGGGCAT GGTGGCAGAT 3000
 GCTGGTAATC TCAGCTATTG AGGAGGCTGA GGCAGGAGAA TTGCTTGAAC CCGGAGGCG 3060
 45 GAGGTGACG TGAGCCAAGA TTGCACCACT GTACTCCAGC CTGGGTGACA AAGTCAGACT 3120
 CCATCTCCAA AAAAAAAAAA AAAA

Seq ID NO: 34 Protein sequence:
 Protein Accession #: NP_055146

50
 1 11 21 31 41 51
 | | | | |
 55 MVRKPVVSTI SKGGYLGQNV NGRLPSLGNK EPPGQEKVQL KRKVLLRGV SIIIGTIIGA 60
 GIFISPKGVL QNTGSVGMSL TIWTVCGVLS LPGALSYABL GTTIKSGGH YTYILEVFGP 120
 LPAFVRVWVE LLIRPAATA VISLAFGRYI LEFFFIQCEI PELAIKLITA VGITVVMVLN 180
 SMSVSNRSRI QIFLTFCKLT AILIIIVPGV MQLIKGQTON FKDAFSGRDS SITRLPLAFY 240
 YGMYAYAGNF YLNFVTEEV NPEKTIPLAI CISMALVTIG YVLTMVAYFT TIMABELLS 300
 60 NAVAVTFSEB LLGNFSLAVP IFVALSCFSG MNGGVFAVR LPYVASREGH LPEILSMIHV 360
 RKHTPLPAVI VLHPLTMIML FSGDLSLLN FLSPARMLPI GLAVAGLIYL RYKCPDMHRP 420
 FKVFLFIPAL FSFTCLFVVA LSLSYDPFST GIGFVITLIG VPAYYLFIIW DKKPRWFRIM 480
 SEKITRTLIQI ILEVVPEEDK L

65 Seq ID NO: 35 Nucleotide sequence:
 Nucleic Acid Accession #: NM_002422
 Coding sequence: 64..1497 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 | | | | |
 ACAAGGAGGC AGGCAAGACA GCAAGGCATA GAGACAACAT AGAGCTAAGT AAAGCCAGTG 60
 GAAATGAGGA GTCTTCCAAT CCTACTGTTC CTGTGCGTGG CAGTTTGCTC AGCCTATCCA 120
 TTGATGAGAG CTGCAAGGGG TGAGGACACC AGCATGAACC TTGTTTCAGAA ATATCTAGAA 180
 75 AACTACTACG ACCTCAAAAA AGATGTGAAA CAGTTTGTTA GGAGAAAGGA CAGTGCTCCT 240
 GTTGTAAAAA AAATCCGAGA AATGCAGAAG TTCTTGGAT TGGAGGTGAC GGGGAAGCTG 300

5 GACTCCGACA CTCTGGAGGT GATGCGCAAG CCCAGGTGTG GAGTTCCTGA TGTGGTCAC 360
 TTCAGAACCT TTCTGGCAT CCGAAGTGG AGGAAAACCC ACCTTACATA CAGGATTGTG 420
 AATTATACAC CAGATTGGCC AAAAGATGCT GTTGATTCTG CTGTTGAGAA AGCTCTGAAA 480
 GTCTGGGAAG AGGTGACTCC ACTCACATTC TCCAGGCTGT ATGAAGGAGA GGCATGATATA 540
 ATGATCTCTT TTGCAGTTAG AGAACATGGA GACTTTTACC CTTTGTATGG ACCTGGAAAT 600
 GTTTTGGCCC ATGCCATAGC CCTGGGCCA GGGATTAAAT GAGATGCCCA CTTTGTATGAT 660
 GATGAACAAAT GGACAAAGGA TACAACAGGG ACCAATTAT TTCTGGTTCG TGCTCATGAA 720
 ATTGGCCACT COCTGGGTCT CTTTCACTCA GCCAACACTG AAGCTTTGAT GTACCCACTC 780
 10 TATCACTCAC TCACAGACCT GACTGGGTTC CGCTGTCTC AAGATGATAT AAATGGCAAT 840
 CAGTCCCTCT ATGGACCTCC CCTGACTCC CCGAGACCC CCTGGTACC CACGGAACCT 900
 GTCCCTCCAG AACCTGGAG GCCAGCCAAC TGTGATCCTG CTTTGTCTCT TGATGCTGTC 960
 AGCACTCTGA GGGGAGAAAT CCTGATCTTT AAAGACAGGC ACTTTTGGCG CAAATCCCTC 1020
 AGGAAGCTTG AACCTGAAT GCATTTGATC TCCTCATTTT GGCCATCTCT TCCTTCAGGC 1080
 GTGGATGCGG CATATGAAGT TACTAGCAAG GACCTCGTTT TCATTTTAA AGGAATCAA 1140
 15 TTCTGGGACA TCAGAGGAAA TGAGGTACGA GCTGGATACC CAAGAGGCAT CCACACCTTA 1200
 GGTTCCTCCT CAACCTTGAG GAAATCGAT GCAGCCATT CTGATAAGGA AAAGAACAAA 1260
 ACATATTCTT TTGTAGAGGA CAAATCTGG AGATTGTATG AGAAGAGAAA TTCCATGGAG 1320
 CCAGCTTTTC CCAAGCAAT AGCTGAAGAC TTTCAGGGA TTGACTCAA GATTGATGCT 1380
 GTTTTGAAG AATTGGGT CTTTATTTT TTTACTGGAT CTTACAGTT GGAGTTTGAC 1440
 20 CCNAATGCAA AGAAGTGAC ACACACTTTG AAGAGTAACA GCTGGCTTAA TTGTGAAG 1500
 AGATATGTAG AAGGCACAAT ATGGCACTT TAAATGAAG TAATAATTCT TCACCTAAGT 1560
 CTCGTGAAT TGAAATGTTT GTTTCTCCT GCCTGTGCTG TGACTCGAGT CACACTCAAG 1620
 GGAACTTGAG CGTGAATCTG TATCTGCGG GTCAATTTTA TGTATTACA GGCATTCAA 1680
 ATGGCTGCT GTTGTAGCTG CACCTGTGCA CATAGAGTGA TCTTCCCAA GAGAAGGGGA 1740
 25 AGCACTCGTG TGCAACAGAC AAGTGACTGT ATCTGTGTAG ACTATTGCT TATTATATAA 1800
 AGACGATTG TCAGTTGTTT T

Seq ID NO: 36 Protein sequence:
 Protein Accession #: NP_002413

35 1 11 21 31 41 51
 MKSLPILLLL CVAVCSAYPL DGAARGEDTS MHLVQKYLEN YYDLKDDVQK FVRRKDSGFPV 60
 VKKIREMKQK LGLEVTGKLD SDTLEVMRKP RCGVPDVGHF RTFPGIPKWR KTHLYTRIVN 120
 YTPDLKDAV DSAVEKALKV WEEVTPLTFS RLYEGEADIM ISFAVREHGD FYFPDGPNGV 180
 LAHAYAPRPG INGDAHDDDD EQWTKDTGT NLFLVAHEI GHSLGLFPHA NTEALMYPLY 240
 HSLTDLTRFR LSQDDINGIQ SLVGPFPDPS BTPLVPTBPV PPERCTPANC DPALSFDAVS 300
 40 TLRGELLIFK DRHFWKSLR KLEPBLHLIS SFWPSLPSGV DAAEVTSKD LVFIKQNP 360
 WAIRGNVRA GYPRGIHTLG PPPTVRKIDA ALSDKEKNKT YPFVBDKIWR PDEKRNMBP 420
 GPPKQIABDF PGIDSKIDAV PEEFGFFYFF TGSSQLSEDFP NAKKVTHLTK SNSWLNC

45 Seq ID NO: 37 Nucleotide sequence:
 Nucleic Acid Accession #: NM_003246
 Coding sequence: 112..3624 (underlined sequences correspond to start and stop codons)

50 1 11 21 31 41 51
 GGACGCACAG GCATTCGCCG CGCCCTCCA GCCCTCGCGG CCTCGCCAC CGCTCCCGGC 60
 CGCCGCGCTC CGGTACACAC AGGATCCCTG CTGGGCACCA ACAGCTCCAC CATGGGGCTG 120
 55 GCCTGGGGAC TAGGCGTCTT GTTCTGATG CATGTGTGTG GCACCAACCG CATTCACAGG 180
 YCTGGCGGAG ACAACAGCGT GTTTGACATC TTTGAACCTA CGGGGGCCCG CGCAAGGGG 240
 TCTGGCGCGC GACTGGTGA GGGCCCGGAC CCTTCCAGCC CAGCTTTCCG CATCGAGGAT 300
 GCCAACCTGA TCCCCCTGT GCGTGATGAC AAGTTCCAAG ACCTGGTGA TGCTGTGCGG 360
 GCAGAAAAGG GTTCTCTCT TCTGGCATCC CTGAGGCAGA TGAAGAAGC CGGGGCCACG 420
 60 CTGCTGGCCC TGGAGCGGAA AGACCACTCT GGCAGGTCT CTAGCGTGT GTCCATATGG 480
 AAGCGCGGCA CCTGGACCT CAGCTGACC GTCCAAGGAA AGCAGCACGT GGTGTCTGTG 540
 GAAGAAGCTC TCTGGCAAC CGGCCAGTGG AAGAGCATCA CCTGTTTGT GCAGGAAGAC 600
 AGGGCCGAGC TGTACATCGA CTGTGAAAAG ATGGAGAATG CTGAGTTGA CGTCCCCATC 660
 CAAGCGCTCT TCACCAGAGA CCTGGCCAGC ATCGCCAGAC TCGCATCGC AAAGGGGGGC 720
 65 GTCAATGACA ATTTCCAGG GGTGCTGAG AATGTGAGGT TTGTCTTTGG AACCACACCA 780
 GAAGACATCT TCAGGAACAA AGGCTGTCTC AGCTCTACCA GTGCTCTCT CACCTTGAC 840
 AACACAGTGG TGAATGGTTC CAGCCCTGCC ATCCGCACTA ACTACATTGG CCACAAGACA 900
 AAGGACTTGC AAGCATCTG CGGCATCTCC TGTGATGAGC TGTCCAGCAT GGTCTTGAA 960
 70 CTCAGGGGCC TCGCACCAT TGTGACCAG CTGCAGGACA GCATCGCAA AGTGACTGAA 1020
 GAGAACAAAG AGTTGGCCAA TGAGCTGAGG CGGCTCCCC TATGCTATCA CAACGGAGTT 1080
 CAGTACAGAA ATAAAGAGGA ATGGACTGTT GATAGCTGCA CTGAGTGCA CTGTGAGAAC 1140
 TCACTTACCA TCTGCAAAA GGTGCTCTGC CCTCATGCG CCGTCTCCAA TGCCACAGTT 1200
 CCTGATGAGG AATGCTGTCC TCGCTGTGG CCGACGACT CTGCGGACGA TGGCTGTCT 1260
 75 CCAATGTCCG AGTGGACCTC CTGTTCTACG AGCTGTGGA ATGGAATTCA CGAGCGCGG 1320
 CGCTCTGCG ATAGCTCAA CAACCGATGT GAGGGCTCT CGGTCCAGAC ACGGACCTGC 1380
 CACATTCAGG AGTGTGCAA AAGATTAAA CAGGATGGT GCTGGAGCCA CTGCTCCCG 1440
 TGGTCATCTT GTTCTGTGAC ATGTGTGAT GGTGTGATCA CAAGGATCCG GCTCTGCAAC 1500

5 TCTCCAGCC CCCAGATGAA TGGGAAACCC TGTGAAGGCG AAGCGCGGGA GACCAAGCC 1560
 TGCAGAAAG ACGCTGCCCC CATCAATGGA GGCTGGGGTC CTGTGTCCAC ATGGGACATC 1620
 TGTTCGTGCA CCGTGTGAGG AGGGGTACAG AAACSTAGTC GTCTCTGCAA CAACCCGCA 1680
 CCCAGTTTG GAGGCAAGGA CTGCGTTGGT GATGTAACAG AAACCCAGAT CTGCAACAAG 1740
 CAGGACTGTC CAATTGATGG ATGCCGTGTC AATCCCTGCT TTGCCGCGGT GAAGTGTACT 1800
 AGCTACCCCTG ATGGCAGCTG GAAATGTGGT GCTTGTCCCC CTGGTTACAG TGGAAATGGC 1860
 ATCCAGTGCA CAGATGTTGA TGAGTGCAAA GAAGTGCCCTG ATGCCGTGCTT CAACCACAAT 1920
 GGAGAGCAAC GGTGTGAGAA CAGGACCCCG GGCTACAAC CTCTGCCCTG CCCCCCAGC 1980
 10 TTCAACGCGT CACAGCCCTT CGGCCAGGGT GTGCAACATG CCACGCGCAA CAACAGGTG 2040
 TGCAGGCCCC GTAACCCCTG CACGGATGGG ACCCAAGACT GCAACAAGAA CGCCAAGTGC 2100
 AACTACCTGG GCCACTATAG CGACCCCATG TACCGCTGGG AGTGCAAGCC TGGCTACGCT 2160
 GGCAATGGCA TCATCTGCGG GGAGGACACA GACCTGGATG GCTGGCCCAA TGAGAACCTG 2220
 GTGTGCGTGG CCAATGCGAC TTACCACTGC AAAAAGGATA ATTGCCCAA CCTTCCCAAC 2280
 TCAGGCGGAG AAGACTATGA CAAGGATGGA ATTGGTATG CTGTGATGA TGACGATGAC 2340
 15 AATGATAAAA TTCCAGATGA CAGGACAAC TGTCCATTCC ATTACAACC AGCTCACTAT 2400
 GACTATGACA GAGATGATGT GGGAGACCCG TGTGCAACT GTCCCTACAA CCACCAACCA 2460
 GATCAGGCAG ACACAGACAA CAATGGGGAA GGAGAGCCCT GTGCTGCAGA CATTGATGGA 2520
 GACGCTATCC TCATGAAACG GGCAACTGCG CAGTACGTCT ACAATGTGGA CCAGAGAGAC 2580
 20 ACTGATATGG ATGGGGTTGG AGATCACTGT GACAATTGCC CCTTGGAAAC CAATCCGGAT 2640
 CAGCTGGAAT CTGACTCAGA CGCAATTGGA GATACCTGTG ACACAATCA GGATATTGAT 2700
 GAGATATGCC ACCAGAACAA TCTGGACAA TGTCCTATG TGCCCAATGC CAACCAAGCT 2760
 GACCATGACA AAGATGGCAA GGGAGATGCC TGTGACCAG ATGATGACAA CGATGSCATT 2820
 CCTGATGACA AGGCAACTG CAGACTGTG GCCAATCCCG ACCAGAAGGA CTCTGAAGCG 2880
 25 GATGCTGAGG GTGATGCCCTG CAAAGATGAT TTTGACCATG ACAGTGTGCC AGACATCGAT 2940
 GACATCTGTC CTGAGAATGT TGACATCACT GAGACCGATT TCCGCGGATT CCAGATGATT 3000
 CCTCTGGACC CCAAGGGGAC ATCCCAAAAT GACCTTAAT GGGTTGTAGC CCATCAGGCT 3060
 AAGAATCTCG TCCAGACTGT CAACCTGTAT CCGGACTCG CTGTAGGTTA TGATGAGTTT 3120
 AATGCTGTGG ACTCTAGTGG CACCTTCTTC ATCAACACCG AAAGGGACGA TGACTATGCT 3180
 30 GGAATTTGCT TTGGCTACCA GTCCAGCAGC CGCTTTTATG TTGTGATGTG GAAGCAAGTC 3240
 ACCGACTCCT ACTGGACAC CAACCCACG AGGGCTCAGG GATACCTCGG CCTTCTGTG 3300
 AAGATTGTAA ACTCCACCAC AGGGCCTGGC GAGCACCTGC GGAACGCCCT GTGGCACACA 3360
 GGAACACCCC TTGCCAGGT GCGCACCCCTG TGGCATGACC CTGTCACAT AGGCTGGAAA 3420
 GATTTACCCG CCTACAGATG GCGTCTCAGC CACAGGCCAA AGACGGGTTT CATTAGAGTG 3480
 35 GTGATGATG AAGGGAAGAA AATCATGGCT GACTCAGGAC CCATCTATGA TAAACCTAT 3540
 GCTGTGTGTA GACTAGGTTT GTTTGTCTTC TCTCAAGAAA TGGTGTCTT CTCTGACCTG 3600
 AATATCGAAT GTAGAGATCC CTATATCA AATGTTGAT TGAAAGACTG ATCATAAACC 3660
 AATGCTGTGA TTGCACTTC TGGAACTATG GGCTTGAGAA AACCCCGAG ATCACTTCTC 3720
 CTGTGCTTCC TTCTTTCTG TGCTTGCACT AGTGTGGACT CCTAGAACGT GCGACCTGCC 3780
 40 TCAAGAAAT CAGATTCTCA AAAACAGACT CATCAGCATT CAGCTCCAA TGAATAAGAC 3840
 ATCTTCCAG CATATAACA ATTGCTTTGG TTCTCTTTG AAAAAGCATC TACTTGTCTT 3900
 AGTTGGGAAG GTGCCATTG CACTCTGCTT TTGTCACAGA GCAGGTGCTT ATTGTGAGG 3960
 CATCTCTGAG CAGTGGACTC AAAAGCATTT TCAGGCATGT CAGAGAAGGG AGGACTCACT 4020
 ARAATTAGCA AACAAACCA CCCTGACATC CTCCTTCAGG AACACGGGA GCAGAGGCCA 4080
 45 AAGCACTAAG GGGAGGGGCG ATACCGAGA CGATTGTATG AAGAAAATAT GGAGGAACCTG 4140
 TTACATGTTT GGTACTAAGT CATTTTCAGG GGATTGAAAG ACTATTGCTG GATTTCATGA 4200
 AATGCTGTGA TTGCACTTC TGGAACTATG GGCTTGAGAA AACCCCGAG ATCACTTCTC 4260
 GGAACAAGG ACTGGCTTCT GGAATTCCTC CTGATCCCC ACCCTTACTC ATCACTTCTC 4320
 AGTGGCCAGA ATTAGGGAAT CAGAATCAAA CCAAGTGAAG GCAAGTCTGG CTGCCATTGC 4380
 50 CTGGTCACAT TGAATTTGGT GGCTTCATTC TAGATGTAGC TTGTGCAGAT GTAGCAGGAA 4440
 AATAGGAAA CCTACCATCT CAGTGAGCAC CAGCTGCCCT CCAAGGAGG GGCAGCGGTG 4500
 CTTATATTTT TATGTTTACA ATGCCACAAA ATTATTATCA ACCTAATCAA AACATTCTCT 4560
 TTCTCTTTTT TCCGTAATTA CTAGGTAGTT TTCTAATCT CTCTTTTGGG AGTATGATT 4620
 55 TTTTAAAGTC TTTACGATGT AAAATATTTA TTTTCTACTT ATTCTGGAAG ATCTGGCTGA 4680
 AGGATTATTC ATGGAACAGG AAGAAGCGTA AAGACTATCC ATGTCTATCT TGTGTAGAGT 4740
 CTTCTGACT GTAAAGATTGT AATACAGAT TATTATTAA CTCTGTTCTG CCTGGAAATT 4800
 TAGGCTTATC ACGGAAAGTG TTTGAGAGCA AGTAGTTGAC ATTTATCAGC AAATCTCTTG 4860
 CAAGAACAGC ACAAGGAAAA TCAGTCTAAT AGCTGCTCT GCCCTTGTG CTCAGAGTGG 4920
 ATGTTATGGG ATTCTTTTTT TCTCTGTTT ATCTTTTCAA GTGGAATTAG TTGGTTATCC 4980
 60 ATTGTCAAT GTTTTAAAT GCAAGGAAAG CCATGAGGTC TTCAATACTG TTTTACCCCA 5040
 TCCCTTGTGC ATATTTCAG GGAAGGAA AGCATATACA CTTTCTCTT TCATTCTTCC 5100
 AAAAGAGAAA AAAATGACAA AAGGTGAAC TTACATACAA ATATTACCTC ATTTGTTGTG 5160
 TGACTGAGTA AAGAATTTT GGATCAAGCG GAAAGAGTTT AAGTGTCTAA CAACTTAAA 5220
 GCTACTGTAG TACCTAAAAA TCGAGTGTG TACATAGCAT ARAACTCTG CAGAGAAGTA 5280
 65 TTCCCAATAA GGAATAGCA TTGAAATGTT AAATACAATT TCTGAAAGTT ATGTTTTTTT 5340
 TCTATCATCT GGTATACCAT TGCTTTATTT TTATAAATTA TTTTCTCAIT GCCATTGGAA 5400
 TAGAATATTC AGATTGTGTA GATATGCTAT TTAATAAATT TATCAGGAAA TACTGCTGTG 5460
 AGAGTTAGTA TTTCTATTTT TATATAATGT TNGCACACTG AATTGAAGAA TTGTTGGTTT 5520
 70 TTCTTTTTTT TTGTTTTTTT TTTTTTTTTT TTTTTTTTTT CTTTGTACCT CCATTTTTTA 5580
 CTATTGGCCA ATACCTTTTT CTAGGAATGT GCTTTTTTTT GTACACATTT TTATCCATTT 5640
 TACATTCTAA AGCAGTGTA GTTGATATAT ACTGTTTCTT ATGTACAAGG AACACAATA 5700
 AATCATATGG AAATTTATAT TT

Seq ID NO: 38 Protein sequence:
 Protein Accession #: NP_003237

```

1      11      21      31      41      51
|      |      |      |      |      |
MGLAWGLGLV FLMHVCGTNR IPESGGDNSV FDIPELTGAA RKGSQRRLVK GPDPSSPAFR 60
IEDANLIPFV PDDKFDLVD AVRAEKGFL LSLRQMKKT RGTLLALERK DHSQQVFSVV 120
5  SNGKACTGLDL SLTVQGHQHV VSVBELLAT GQNKSTLFPV QEDRAQLYID CEHOMENALD 180
VPIQSVPFTR LASIARLRLA KGGVNDNPGG VLQNVRFVFG TTPEDILRNK GCSSSTSVLL 240
TLDNNVNGS SPARKTNYIG HKTKDLQAIC GISCDLSSM VLELRGLRTI VTTLQDSIRK 300
VTRENKELAN ELRRPPLCYH NGVQYRNNEE WTVDSCTECH QNSVTICKK VSCPIMPCHN 360
ATVPDGECCP RCWPSDSADD GWSFWSWETS CSTSCQNGIQ QGRSICDSLW NRCEGSSVQT 420
10 RTCHIQCCKK RFKQDGGWSH WSPWSSCSVT CGDGVITRIR LCNSPSPQMN GKPCBGEARE 480
TKACKKDACP INGGWGWSPF WDICSVTOGG GVQKRSRLCN NPAPQFGGKD CVGDVTENQI 540
CNQDQCPIDG CLSNPCFAGV KCTSYFDGSW KOGACPPGYS GNGIQCTQVD ECKEVPDADF 600
NHNGEHCEN TDPGYNCLPC PFRFTGSGPF GQGVHEHATAN KQVCKPRNPC TDGTHDCNKN 660
AKCNYLGHYS DPMVRCCKP GYAGNGIICG EDTDLGWFN ENLVCVANAT YHCKKDCNPN 720
15 LPNSGGQEDY KDGIGDACDD DDDNDKIPDD RDNCPFFHYP AQYDYDRDDV GDRCDNCPYN 780
HNPDAQADTN NGEGDACAAD IDGGDILNER DMCQYVNVND QRDITMDGVG DQCDNCPLEH 840
NPDQLDSDS RIGDTCDNNQ DIDEDGHQNN LDNCPYVNA NQADHDKDGK GDACDHDDN 900
DGIPLDKNC RLVPNPQKD SDGDRGDAC KDDFDHDSVP DIDDICPENV DISBTFRRP 960
20 QMPLPFGT SQNDPMVVR HQGKBLVQTV NCDPLAVGY DEFNAVDFSG TFFINTERDD 1020
DYAGVYFQYQ SSSRFYVMW KQVTSYWDY NPTAQGYSG LSVKVNSTT GPGEHLRNAL 1080
WHTGNTPQV RTHLHDPRI GWKDFAYRW RLSHRPKTF IRVVMYBGKK IMADSGPIYD 1140
KTYAGRGLGL FVFSQEMVFF SDLKYECDP

```

Seq ID NO: 39 Nucleotide sequence.

Nucleic Acid Accession #: BC004299

Coding sequence: 69..1235 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
|      |      |      |      |      |
CCCCACCCCT GCGAGGCCA GGTCCGCGCC TCCCCGCCA GCGAAGCGA GCGGACCCGC 60
GTGCGGCCAT GGCTTCGCTG CTGGGAGCCT ACCCTTGGCC CGAGGGTCTC GAGTGCCCGG 120
35 CCTTGAACGC CGAGCTGTCG GATGGACAAT CGCCGCGGCG CGTCCCCGGG CCCCAGGGGG 180
ACAAGGGCTC CGAGAGCCGT ATCCGCGCGG CCATGAACGC CTTTCATGTT TGGGCCAAGG 240
ACGAGAGGAA ACGGCTGGCA GTGCAGAAC CCAGACTGCA CAACGCGGAG CTACAGCAAG 300
TGCTGGGAAA GTGCTGAAG CGCTGACGCG TGTCGCAAG GAGGCGGTAC GTGAGCAGAG 360
CGAGCGGCTC GCGCTGCAG CACATGCAGG ACTACCCCAA CTACAGTAC CGGCGCGGCA 420
40 GGAAGAGCA GGCACAGCGG CTGTGCAAGC GCGTGGACCC GGGCTTCCTT CTGAGCTCCC 480
TCTCCCGGGA CCAGAACGCC CTGCCGAGGA AGAGAAGCGG CAGCGCGGGG GCGCTGGGGG 540
AGAAAGAGGA CAGGGGTGAG TACTCCCGCG CACTCCCTCT GCCCAGCCTC CGGGGCTGCT 600
ACCAAGAGGG GCGGCTGCTG GTTGGCGGCG GCGGCACCCC GAGCAGTGTG GACAGTACC 660
CGTACGGGCT GCCACACCT CCGTGAATGT CTCCTCTGGA CGTCTGGAG CCGGAGCAGA 720
45 CCTTCTTCTC TCCCGCTGCG CAGGAGGAGC ATGGCCATCC CGCGCGCCTC CCCCACCTGC 780
CAGGCGACCC GTACTCACCG GAGTACGCCC CAAGCCCTCT CCACGTAGC CAGCCCTGG 840
GCTCCCTGCG CCTTGGCCAG TCCCGCGGCG TCTCCATGAT GTCCCTCTGA CCGGCTGTC 900
CCCCATCTCC TGCCATTATC TCCCGCGGCA CCTACCAACC ACTCCACTCC AACCTCCAAG 960
CCACCTCGGG CCAGCTTTCC CGGCTCTCTG AGCACCTGCG CTTCGACGCC CTGGATCAAC 1020
50 TGAGCCAGGT GGAACCTCTG GGGACATGG ATCGCAATGA ATTCCACCAG TATTTGAACA 1080
CTCTCGGCCA CCCAGACTCC GCCACAGGGG CCATGGCCCTC CAGTGGGCGT GTTCCGGTCT 1140
CCCAGGTGAC ACCAAGCGGT CCCACAGAGA CCAGCCTCAT CTCCTCTCTG GCTGATGCCA 1200
CGGCCACGTA CTACAACAGC TACAGTGTGT CATAGAGCTG GAGGCGCCCC CTCCGCTCAG 1260
CCCTCGCGCC CTCTCCTTCT TGTGCCTTGA GTGGCAGAGG AGCCGCTCCG CCACACAGC 1320
55 TTTCTCTCCA CGCTCAGGG CAGGGAGGTC TGAACCTGCG CCCCAGAGCC TTTGGCCTAA 1380
GCTGAGCTCT CCTTATCCGA GTGCCGCTCT TATCCCTTCC CCCAGTTCC AGCCCCTGCA 1440
GCCACATTTT TAAGTATATT CCTTCAAGTG AGTTTCTCTC CAGCCCTGGA GAGTTGCTGT 1500
CTCCAGTGGG AATGTTCACT GACGCTTTTT CTGTGAGGCC ATCATCGAAA CTAATGGGGG 1560
GACAGACTTG ATAGCCAAGG TCCCTTCTGG TCCAGTTTTC TGATTTAGGG TTCTCTCAAG 1620
60 ATTAATAAAG GAAGATGGGG AAATTTGACT CATTAAAGAG CTCGCTAACC TACGATCTGG 1680
TGATAATTTT GTGTGCACAG CCCAAGGACC ACGAGGCTTT CTGCACTTTC TGCAACCCCT 1740
TCCAAGTGA CCACAAATTT TCAAAGGGAC TCATACAAAT TGAGAAAAAA CAGTCAACCT 1800
GATTTGAGAA ATTAACCAAT ATGGCTAACT ATATCAGAGA AAATGGGATT GAGTTAAAC 1860
TATTTTATTT TAAATATACA TTTTAAAGCA GTTCTTTTTT TTTGTTAATT TGTTTATTAT 1920
65 ACACACACTT CAAGAGCCAC CGCGCCCGAG CTACATTATAT AATTTTCATT CTCTTTTACC 1980
TATAAAATTC AGTGTATTAG TTTCAATTACA TAGGAGAAAT TATATTTCTA AACATTTTAT 2040
GATGTTTAAA AACAAAACAG GCTGTTGTAA AAAAAAATA AAAAAAATA

```

Seq ID NO: 40 Protein sequence.

Protein Accession #: AAH04299

```

1      11      21      31      41      51
|      |      |      |      |      |
MASLLGAYPW PEGLECPALD AELSDGQSPF AVPRFPQDKG SESRIRRPMM AFMVWAKDER 60
KRLAVNQDL HNALSRLMLG KSWKALTLGQ KRFPVDEAER LRLQHMODYP NYKYRPRRKK 120
75 QAKRLCKRVD PGFLLSLSLR DQNALPEKRS GSRGALGEKE DRGEYSPGTA LPRLRGCTHE 180

```

GPAGGGGGGT PSSVDITYFYG LPTPPMSPL DVLEPEQTFF SSPQBEHGH PRRIPLPGH 240
PYSFEPYAPSP LHCGRPLGSL ALGQSPGVSM MSPVPGCPPS PAYYSPATYH PLHSNLQHL 300
GQLSPPPPEHP GFDALDQLSQ VELLGDMDRN EPDQYLAIFG HPDSATGAMA LSGHVVPVSQV 360
TFTGTPTETSL ISVLADATAT YNYSYSVS

5

Seq ID NO: 41 Nucleotide sequence:

Nucleic Acid Accession #: NM_004449

10 Coding sequence: 1..1389 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
15 ATGATTGAGA CTGTCCCGGA CCCAGCAGCT CATATCAAGG AAGCCTTATC AGTTGTGAGT 60
GAGGACCACT CGTTGTTTGA GTGTGCCTAC GGAACGCCAC ACCTGGCTAA GACAGAGATG 120
ACCGGCTCCT CCTCCAGCGA CTATGGACAG ACTTCCAAGA TGAGCCCAAG CGTCCCTCAG 180
CAGGATTGGC TGCTCAACCC CCCAGCCAGG GTCAACCATCA AAATGGAATG TAACCTTAGC 240
CAGGTGAATG GCTCAAGGAA CTCTCCTGAT GAATGCAGTG TGGCCAAAGG CGGGAAGATG 300
20 GTGGGCAGCC CAGACACCGT TGGGATGAAC TACGGCAGCT ACATGGAGGA GAAGCACAATG 360
CCACCCCAAA ACATGACCAC GAACGAGGCG AGAGTTATCG TGCCAGCAGA TCCTACGCTA 420
TGAGGTACAG ACCATGTGCG GCAGTGGCTG GAGTGGGCGG TGAAAGAATA TGGCCTTCCA 480
GAGCTCAACA TCTTGTTATT CCAGAACATC GATGGGAAGG AACTGTGCAA GATGACCAG 540
GAGGACTTCC AGAGGCTCAC CCCAGCTAC AACGCGGACA TCCTTCTCTC ACATCTCCAC 600
25 TACCTCAGAG AGACTCCTCT TCCACATTG ACTTCAGATG ATGTTGATAA AGCCTTACAA 660
AACTCTCCAC GGTTAATGCA TGCTAGAAAC ACAGATTAC CATATGAGCC CCCAGGAGA 720
TCAGCCTGGA CCGGTACCGG CCACCCACG CCCAGTCCA AAGCTGCTCA ACCATCTCCT 780
TCCACAGTGC CCRAAACTGA AGACCAGGT CCTCAGTTAG ATCCTTATCA GATTCTTGA 840
CCACCAAGTA CCGCCTTGC AAATCCAGGC AGTGGCCAGA TCAGCTTTG GCAGTTCCTC 900
30 CTGAGCTTCC TGTCGACAG CTCCAACTCC AGCTGCATCA CCTGGGAAGG CACCAACGGG 960
GAGTTCAGTA TGACGGATCC GCACGAGGTG GCCCGGCCCT GGGGAGAGCG GAAGAGCAAA 1020
CCCAACATGA ACTACGATAA GCTCAGCCCG GCCCTCCGTT ACTACTATGA CAAGAACATC 1080
ATGACCAAGG TCCATGGGAA GCGCTACGCC TACAAGTTCG ACTTCCACGG GATCGCCAG 1140
GCCCTCCAGC CCCACCCGCC GGAGTCATCT CTGTACAAGT ACCCTCAGA CCTCCCGTAC 1200
35 ATGGGCTCCT ATCACGCCCA CCCACAGAAG ATGAACCTTG TGGCGCCCA CCTCCAGCC 1260
CTCCCCGTGA CATCTTCCAG TTTTTTTGCT GCCCAAAACC CATACTGGAA TTCACCAACT 1320
GGGGGTATAT ACCCCAACAC TAGGCTCCTCC ACCAGCCATA TGCTTCTCA TCTGGGCACT 1380
TACTACTAA

40

Seq ID NO: 42 Protein sequence:

Protein Accession #: NP_004440

1 11 21 31 41 51
45 MIQTVDPFAA HIKEALSVVS EDQSLFECAY GTPHLAKTEM TASSSDYDYG TSXMSPRVPQ 60
QDWLQSPPAR VTIKMRCNPS QVNGSRNSPD ECSVAKGGKM VGSPTDTVMN YGSYMEKHM 120
50 PPMNMTNER RVIVPADPTL WSTDHVRQWL EWAUVKEYGLP DVNILLPQNI DGKELCKMTK 180
DDPQLRTPSY NADILLSHLH YLRETPPLPHL TSDDVDKALQ NSPRLMHARN TDLFVEPPRR 240
SAWTHGHRPT PQSKAAQSPS STVPKTEDQR PQLDPYQILG PTSSRLANPG SQQIQLWQFL 300
LELLSDSSNS SCITWGTNG EFKMTDPDEV ARRWGRKSK PMNNDKLSR ALRYYYDKNI 360
MTKVHGKRYA YKFPFHGLAQ ALQPHPPSS LYKYPDDLPH MGSYHAHPQK MNFVAPHPPA 420
55 LPVTSSSFFA APNFWNSPT GGIYPNTRLP TSHMPSHLGT YY

Seq ID NO: 43 Nucleotide sequence:

Nucleic Acid Accession #: NM_005100

60 Coding sequence: 192..5537 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
65 CCTTCTTTTA AGGAGTTTGC CGCGAGCGCG TCTCCTTCAT TCGAGGCTG GCGCGTTCG 60
CAGTGGGCTG GCGGCGAAGG AAGGCGCTCT CGGAGCTCA CGGCGCGCG TCTTTTGGCT 120
CTTGGCCCTG TCCCTGCGGC TTGGGAAAG CGTAACCGG CGGCTAGCG CGGAGAAATG 180
GCGGAGAGC CTTGGGCGCC GGGAGCTCCA CCGAGCAGCG CAGCCCGAG CAGCGGCCG 240
AGGGGAGCTC CAGCCCGGCT GAGCCCGAGC CCAGCGCGCG CGGCCCTCG GCGGAGCGG 300
CGCCAGACAC CACCGCGGAC CCGCCATCG CTGCTCGGA CCGGCCACC AAGCTCCTAC 360
70 AGAAGAATGG TCAGCTGTCC ACCATCAATG GCGTAGCTGA GCAAGATGAG CTCAGCCTCC 420
AAGAGGOTGA CCTAAATGGC CAGAAAGGAG CCCTGAACGG TCAAGGAGCC CTAAACAGCC 480
AGGAGGAAGA AGAAGTCATT GTACGAGAGG TTGGACAGAG AGACTCTGAA GATGTGAGCG 540
AAGAGAGACT CGATAAAGAG ATGGCTACTA AGTCAGCGGT TGTTCAGAC ATCAGAGATG 600
ATGGGCGAGA GGAGAAACCGA AATATCGAAC AGATTCCTTC TTCAGAAAGC AATTAGAAAG 660
75 AGCTTAACACA ACCCACTGAG TCCAGGCTA ATGATATTGG ATTTAAGAAG GTGTTTAAGT 720
TTGTGGCTT TAAATTCACT GTGAAAAGG ATAAGACAGA GAAGCTGAC ACTGTCCAGC 780

	TACTCACTGT	GAAGAAAGAT	GAAGGGGAGG	GAGCAGCAGG	GGCTGGCGAC	CACCAGGACC	840
	CCAGCCTTGG	GGCTGGAGAA	GCAGCATCCA	AAGAAAGGGA	ACCCAAACAA	TCTACAGAGA	900
	AACCCGAGAA	GACCTTGAA	OGTGAGCAAA	GCCACGCAGA	AATTTCTCCC	CCAGCCGAAT	960
	CTGGCAGCAG	AGTGGAGGAA	TGCAAGAGAG	AAGGAGAGAA	GAACCAAGAA	AAAGAACCTA	1020
5	GCAAGTCTGC	AGAATCTCCG	ACTAGTCCCG	TGACCAAGTA	AACAGGATCA	ACCTTCAAAA	1080
	AATTTCTTAC	TCAAGATTGG	GCCCGCTGGC	GCAAAAAGAC	CAGTTTTCAG	AAGCCGAAGG	1140
	AGGATGAAGT	GGAAAGCTTC	GAGAAAGAAA	AGGAACAAGA	GCCAGAAAAA	GTAGACACAG	1200
	AAGAAGACGG	AAAGGCAGAG	GTTCCTCCCG	AGAAACTGAC	GCCCTCCGAG	GAGGCCACCC	1260
	CACAGGAGCC	GGCAGAAAGT	GCCCACGAGC	CCCGGTTATC	AGCTGAATAT	GAGAAAGTTG	1320
10	AGCTGCGCTC	AGAGGAGCAA	GTCACTGGCT	CGCAGGGACC	TTCTGAAGAG	AAACCTGCTC	1380
	CGTTGGCGAC	AGAAAGTGT	GATGAGAAAA	TAGAAGTCCA	CCAAGAAAG	GTGTGGCCCG	1440
	AAGTCCAGCT	CAGCACCGTG	GAGGAGAGAA	CCGAAAGAGC	GAACACGGAG	GTGGAAGAAA	1500
	CAGCAGGCTC	TGTGCCAGCT	GAAGAATTGG	TTGGAATGGA	TGCAGAACCT	CAGGAAGCCG	1560
	AACCTGCGAA	GGAGCTGGTG	AAGCTCAAG	AAACGTGTGT	TTCCGGAGAG	GACCTTACAC	1620
15	AGGAGAGTGA	CCTCAGTCTC	GATGAGAAAG	TGCTGTCCAA	ACCCCCCGAA	GGCGTTGTGA	1680
	GTGAGGTGGA	AATGCTGTCA	TCACAGGAGA	GAATGAAGGT	GCAGGGAAGT	CCACTAAAGA	1740
	AGCTTTTATC	CAGCAGTGGC	TTAAAAAAGC	TTTCTGGAAA	GAACAGAGAA	GGGAAAAAG	1800
	GAGGAGGAGA	CGAGGAATCA	GGGGAGCACA	CTCAGGTTCC	AGCGATTTCT	CCGGACAGCC	1860
	AGGAGGAGCA	AAAGGGCGAG	AGCTCTGCTC	CATCCCTTGA	GGAGCCCGAG	GAGATCACGT	1920
20	CTCTGGAAAA	GGGCTTAGCC	GAGGTGCAGC	AGGATGGGGA	AGCTGAAGAA	GGAGTCACTT	1980
	CCGATGGAGA	GAIAAAAAAG	GAAGGTGTCA	CTCCCTGGGC	ATCATTTCAA	AAGATGGTGA	2040
	CGCCAGAGAA	GGGTGTAGAA	CGGCTTCCGG	AAAGTGATAA	AGAAGATGAG	CTGACAAAGG	2100
	TCAGAGCGCC	TACCTTGTCT	TCCACCGAGA	GCACAGCCTC	TGAATGCGAA	GAGAAATGA	2160
	AAGGGAGCGT	GGAAAGGCCA	AAGCCCGAAG	AAACAAAGCG	CAAGTGGGAT	ACCTCAGTAT	2220
25	CTTGGGAAGC	TTTAATTTGT	GTGGGATCAT	CCAAGAAAG	AGCAAGGAGA	AGGTCTCTTT	2280
	CTGATGAGGA	AGGGGGACCA	AAAGCAATGG	GAGGAGACCA	CCAGAAAGCT	GATGAGGCCG	2340
	GAIAAGACAA	AGAGACGGGG	ACAGACGGGA	TCCTTGCTGG	TTCCCAAGAA	CATGATCCAG	2400
	GGCAGGGAAG	TTCTCCCGCG	GAGCAAGCTG	GAAGCCCTAC	CGAAGGGGAG	GGCGTTTCCA	2460
	CCTGGGAGTC	ATTTAAAGGG	TTAGTCAAGC	CAAGAAAAAA	ATCAAGTCCC	AAGCTGGAAG	2520
30	AGAAAGCGGA	AGACTCCATA	GCTGGGTCTG	GTGTAGAACA	TTCCACTCCA	GACACTGAAC	2580
	CGGTAAAGAA	AGAATCTCTG	GTCTCAATCA	AGAAGTTTAT	TCCTGGAAGA	AGGAAGAAAA	2640
	GGCCAGATGG	GAACAAGGAA	CAAGCCCTCG	TTGAAGAAGC	AGGGCCACAA	GGGGCCAAAG	2700
	AAGATGACTC	TGATGTCCCG	GCCGTGGTCC	CTCTGTCTGA	GTATGATGCT	GTAGAAAGGG	2760
	AGAAATATGA	GGCACAGCAA	GCCCAAAAAG	GCGCAGAGCA	GCCCGAGCAG	AAGGCAGCCA	2820
35	CTGAGGTGTC	CAAGGAGCTC	AGCGAGAGTC	AGGTTCATAT	GATGGCAGCA	GCTGTGCTGT	2880
	ACGGGACGAG	GGCAGCTACC	ATTATTGAAG	AAAGCTTCCC	TTCTTGGATA	TCTGTCTCAG	2940
	TGACAGAAAC	TCTTGAACAA	GTAGAAGCTG	AAGCCGCACT	TTTAAGTGA	GAGGTATTGG	3000
	AAAGAGAAGT	AATTGACAGAA	GAAGAACCCC	CCACGGTTAC	TGAACCTCTG	CCAGGAAACA	3060
	GAGAGGCGCC	GGGCGACACG	GTCCGTTAGT	AGGCGGAATT	GACCCCGGAA	GCTGTGACAG	3120
40	CTGCAAGAAC	TGCAAGGCCA	TTGGGTTCGG	AAGAAGGAAC	CGAAGCATCT	GCTGTGGAAG	3180
	AGACACAGCA	AATGGTGTCA	GCAGTCTCCC	AGTTAACCGA	CTCCCCAGAG	ACCAAGAGGG	3240
	AGGCCACTCC	GGTGCAGGAG	GTGGAAGGTG	GGGTACCTGA	CATAGAAGAG	CAAGAGAGGC	3300
	GGACTCAAGA	GGTCTCCAG	GCAGTGGCAG	AAAAAGTGAA	AGAGGAATCC	CAGCTGCGTG	3360
45	GCACCGTGG	GCCAGAAAGT	GTGCTTCAGC	CTGTGCAGAG	AGCAGAGGCA	GAAGAGCCAG	3420
	AAGAGCAGGC	TGAAGCGTCG	GGTCTGAAGA	AAGAGACGGA	TGTAGTGTGG	AAAGTAGATG	3480
	CTCAGGAGCC	AAAAAAGTGG	CCTTTTACAC	AAGGGAAGGT	GGTGGGCGAG	ACCAACCCAG	3540
	AAAGCTTTGA	AAAAGCTCCT	CAAGTCAAG	AGAGCATAGA	GTCCAGTGA	CTTGTAAACA	3600
	CTGTTCAGAG	CGAAACCTTA	GCTGGGGTAA	AATCACAGGA	GATGGTGAAT	GAACAGGCTA	3660
50	TCCGCCCTGA	CTGCTGGGAA	ACCCCTACAG	ACAGTGAAGC	TGATGGAAGC	ACCCCGGTAG	3720
	CCGACTTTGA	CGCACAGGCG	ACAACCCAGA	AAGACGAGAT	TGTGGAAATC	CATGAGGAGA	3780
	ATGAGGTGCG	ATCTGCTACC	CAGTCAGGGG	GCACAGAGCC	AGAGGCAAGT	CCTGCACAGA	3840
	AAGAGAGGCC	TCCAGCACCT	TCCAGTTTGT	TGTTCCAGGA	AGAAACTAAA	GAACAATCAA	3900
	AGATGGAGAA	CATCTAGAG	CATACAGATA	AAGAGGTGTC	AGTGGAAACT	GTATCCATTC	3960
	TGTCAAGAAC	TGAGGGGACT	CAAGAGGCTG	ACCAGTATGC	TGATGAGAAA	ACCAAGAGCG	4020
55	TACCAATTTT	CGAAGGACTT	GAGGGGTCTA	TAGACACAGG	CATACAGTTC	AGTCCGGAAA	4080
	AGGTCACTGA	AGTTGCCCTT	AAAGGTGAAG	GGACAGAGAA	AGCTGAATGT	AAAAAGGATG	4140
	ATGCTCTTGA	ACTGCAGAGT	CACGCTAAGT	CTCCTCCATC	CCCGTGGAG	AGAGAGATGG	4200
	TAGTTCAAGT	CGAAAGGGAG	AAAACAGAA	CAGAGCCAAC	CCATGTGAAT	GAAGAGAGCC	4260
	TTGAGCAGCA	AACAGCTGTT	ACGTATCTGT	AAGAGGTGAG	TAAGCAGCTC	CTCCAGACAG	4320
60	TGAATGTGCC	CATCATAGAT	GGGGCAAGG	AAATCAGCAG	TTTGAAGGAA	AGCCCTCTCT	4380
	CCTGTCTAGG	TCAAGAGGAG	GCAGTATGCA	CCAAAAATCA	AGTTCAAGAG	TCTGAGGCAT	4440
	CATTCACTCT	AACAGCGGCT	GCAGAGGAGG	AAAAGGTCTT	AGGAGAAACT	GCCAACTTTT	4500
	TAGAAACAGG	TGAAACGTTG	GAGCCTGCAG	GTGCACATTT	AGTTCTGGAA	GAGAAATCCT	4560
	CTGAAAAAAA	TGAAGACTTT	GCGCTCATC	CAGGGGAAGA	TGCTGTGCC	ACAGGGCCCG	4620
65	ACTGTCAAGC	AAAATCGACA	CCAGTGAATG	TATCTGTAC	TACCAAGAAA	GGCTTAAGTT	4680
	CCGACTGGGA	AGGAGAGAAA	ACCACATCAC	TGAAGTGGAA	GTCAAGTGA	GTGATGAGC	4740
	AGGTGTCTTG	CCAGGAGGTC	AAAGTGAAGT	TAGCAATTGA	GGATTTAGAG	CCTGAAATG	4800
	GGATTTGTGA	ACTTGAGACC	AAAAGCAGTA	AACTTGTCCA	AAACATCATC	CAGACAGCCG	4860
	TTGACCAAGT	TGTACGTACA	GAAGAAACAG	CCACCGAAAT	GTGACCTCT	GAGTTACAGA	4920
70	CACAGCTCA	CGTGATAAAA	GCTGACAGCC	AGGACGCTGG	ACAGGAAACG	GAGAAAGAG	4980
	GAGAGGAAAC	TCAGGCTCTC	GCACAGGATG	AAACACCAAT	TACTTCAGCC	AAAGAGGAGT	5040
	CAGAGTCAAC	CGCAGTGGGA	CAAGCACATT	CTGATATTTT	CAAGACATGT	AGTGAAGCCT	5100
	CAGAAAGAGC	CATGACTGTT	GAGGTAGAA	GTTCACACTG	AAATGATCAG	CAGCTGGAAG	5160
	AGGTGCTCT	CCCATCTGAG	GAAGAGGGAG	GTGGAGCTGG	AAACAACTCT	GTGCCAGAG	5220
75	TGCTGTGCTA	TGCTTGTGA	GCAGAAAGAA	TAGAGAAAGT	ACTAGTTGAA	CGAAAGAGAG	5280
	ATGAAAAAGG	TGATGATGTT	GATGACCCCT	AAAACAGAAA	CTCAGCCCTG	GCTGATACCT	5340

5 ATGCTCTCAGG AGGCTTAACC AAAGAGTCCC CAGATACAAA TGGACCAAAA CAAAAAGAGA 5400
 AGGAGGATGC CCAGGAAGTA GAATTGCAGG AAGGAAAAGT GCACAGTGAA TCAGATAAAG 5460
 CGATCACACC CCAAGCACAG GAGGAGTTAC AGAAACAAGA GAGAGAATCT GCAGAGTCAG 5520
 AACTTACAGA ATCTTAAAC ATCATGCACT TAAACTCATT GTCTGTTTGG AAGACCAGAA 5580
 TGTGAAGACA AGTAGTAGAA GAAAATGAAT GCTGCTGCTG AGACTGAAGA CCAGTATTTC 5640
 AGAAGCTTGA GAATTGGAGA GCAGGCACAT CAAGTATCTT CATTCTAGA GAGCCCTGA 5700
 CAATCTCTAG GCTTCATCAG GAGCTAGAGC CATTTAACAT TTCTCTTTTC CAAGACCAAC 5760
 CTACAAATTT CCCTTGATAA CCATATAAAT TCTGATTAA GTCTCTAAAT TCTTAACCTG 5820
 GAACTGGATG TGGCAATACC TAGTCTGCT TCTGAAACTG GAGTATCATT CTTTACATAT 5880
 10 TTATATGTAT GTTTTAAGTA GTCTCTCTGT ATCTATTGTA TATTTTTTTC TTAATGTTTA 5940
 AGGAAATGTG CAGGATACTA CATGCTTTT GTATCACACA GTATATGATG GGGCATGTGC 6000
 CATAGTGCAG GCTTGGGGAG CTTTAAGCCT CAGTTATATA ACCCACAAA AACAGAGCCT 6060
 CCTAGATGTA ACATTCTCTGA TCAAGGTACA ATTCTTTAAA ATTCACATAAT GATTGAGGTC 6120
 CATATTAGT GGTACTCTGA AATTGGTCAC TTCTCTATTA CAGGAGTGT GCCAAAACATA 6180
 15 AAAAGCATT TGAACATAC AGAATGTTCT ATTGTATTG GGAATTTTG CTTTCTAACC 6240
 CAGTGGAGGT TAGAAGAAG TTATATTCTG GTAGCAAAAT AACTTTACAT CTTTCTCTCT 6300
 ACTTGTATG GTTGTGTTGA CCGATAAGTG TGCTTAATCC TGAGGCAAG TAGTGAATAT 6360
 GTTTTATATG TTATGAAGAA AAGAATTGTT GTAAGTTTT GATTCTACTT TTATATGCTG 6420
 GACTGCAATC ACACATGGCA TGAATAAGT CAGGTCTCTT ACAATGTTA TTTTGTATGA 6480
 20 TACTGGATTG TGTGTTGCC ATATTGTGTC CATTCTTTA AGAACAATGT TGCAACACAT 6540
 TCATTGGAT AAGTTGTGAT TTAGCAGCTG ATTTAAATAA AATATTGTCT TCCTTAAAA 6600
 AAAAAAA

Seq ID NO: 44 Protein sequence:
 Protein Accession #: NP_005091

30 1 11 21 31 41 51
 MGAGSSTQR SPEQPPEGSS TPAEPEPSGG GPSAEAAPDT TADPAIAASD PATKLLQKNG 60
 QLSTINGVAE QDELSLQEGD LMGQKGAING QGALNSQHEE EVIVTEVQQR DSEDVSRDS 120
 DKENATKSAV VHDITDDGQG ENRNIEQIPS SESNLEELTQ PTESQANDIG PKKVFKFVGF 180
 KPTVKDKTKE KPDVTQLLTV KKDEGEAAG AGDHQDPSLG AGEAASKESB PKQSTKPRE 240
 35 TLKREGSHAE ISPPAESGGQ VBECKEGBEE KQEKESPSKSA ESPTSPVTSB TGSTFKKFFT 300
 QGWAGWRKKT SPRPKEDSV EASEKKKEQE PEKVDTEEDG KAEVASEKLT ASEQAHQPBP 360
 AESHAEPRLS AEYKVELPS BEQVSSSQGP SEBKPAPLAT EVFDEKIEVH QEEVVAHVH 420
 STVEERTBEQ KTEVESTAGS VPAEELVGM DAEQEAAPAK ELVKLEKTCV SGEDPTQGD 480
 LSPDEKVLKS PPEGVVSEVE MLSSQERMKV QGSLPKLFT STGLKKLSGK KQKGRGGGD 540
 40 EESGEHTQVP ADSPDSQEQ KRSSASPSB EPERITCLEK GLAEVQDGE AEEGATSDGE 600
 KKRGEVTPWA SPKKNVTPKK RVRRPSESDB EDELDRVESA TLESSTESTAS EMQEMKGSV 660
 EEPKPEBPKR KVDTSVSWEA LICVSSSKKR ARRRSSDSE GPKAMGGDH QKADEAGDK 720
 BTGTDGILAG SQEHDPPQGS SSPAQAGSPT EGEGVSTWES FKRLVTPRKK SKSKLEEKSE 780
 DSIAGSGVHE STPDTEPGKE ESWSIKKFI PGRKKRPDG KQEQAPVEDA GPTGANEDDS 840
 45 DVPAVPLSE YDAVEREKME AQQAQKGAEBQ PEQKAATEVS KELSESQVHM MAADVADGTR 900
 AATTIEERSP SWISASVTEP LEQVEAEAL LTHEVLEREV IAESEPTVT EPLPENREAR 960
 GDTVSEAEAL TPEAVTAEST AGPLGSEEGT EASAAETTE MSAVSQLT D SPDTTEATP 1020
 VQEVGEGVDP IEQERRTQE VLQVAEKVK BESQLPTGG PEDVLQPVQR AEAEERPEQA 1080
 EASGLKKETD VVLKVDAQBA KTEPPTQGV VQOTTPESE KAPQVTESE SSELVTTQQA 1140
 50 ETLAGVKSQE MWMEQAIPPD SVETPTDSET DGSTPVADPD APGTTQKDEI VEIHEBNEVA 1200
 SGTQSGGTEA EAVPAQKERP PAPSSFPVQE ETKEQSKMED TLEHTDKESV VETVSILSKT 1260
 EGTQEQADQYA DEKTKDVPPF EGLEGSIDTG ITVSRKRVTE VALKRGTESE AECKDDDALE 1320
 LQSHAKSPPS PVEREMVQV EREKTEABPT HVNEEKLHE TAVTVSEVS KQLLQTVNVP 1380
 IIDGAKEVSS LEGSPFPCLG QEEAVCTKI QVSSEASFTL TAAAREKVL GETANILETG 1440
 55 ETLBPAGHAL VLEKSSSEKN EDPAHPGED AVPTGPDQCA KSTPVIVSAT TKKGLSDLE 1500
 GEKTTSLKWK SDEVDQVAC QEVKVSVAIE DLEPENGILE LETKSKLVQ NIIQTAVDQF 1560
 VRTEETATEM LTSELQTAH VIKADSQDAG QSTKEGBEP QASAQDETP TSARESEST 1620
 AVGQAHSDIS KMSSEASEKT MTVEVEGSTV NDQQLBEVVL PSEEBGGAG TKSVPEDDGH 1680
 60 ALLAEKIEKS LVEPKEDSKG DDVDDPENQN SALADTDASG GLTKESPDTN GPKQKEKEDA 1740
 QSEVLQGVK HSSEDKAITP QAQELQKQE RESAKSELTE S

Seq ID NO: 45 Nucleotide sequence:
 Nucleic Acid Accession #: NM_001290
 Coding sequence: 110..1231 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51
 GTGAGCGTGT GTGCGTGCCT TACTTTGTA CTGGGAAGAA CACAGCCCAT GTGCTCTGCA 60
 TGGACGTTAC TGATACTCTG TTTAGCTTGA TTTTGGAAAA GCAGGCAAGA TGTCCAGCAC 120
 70 ACCAGATGAC CCTTCTATT CTCTCTCTTT CGGCCCATTT TATAGGAGGC ATACACCATA 180
 CATGCTACAG CCAGAGTACC GAATCTATGA GATGAACAG AGACTGCAGT CTCGCACAGA 240
 GGTATGATAC AACCTCTGTT GGGCAGCCTT TGCCACTGAA TTTTGTGAAG ATGACGCCAC 300
 ATTAACCTCT TCAATTTGTT TGGAGATGG ACCAAGCGA TACACTATCG GCAGGACCTC 360
 75 CATCCCCGCT TACTTTAGCA CTGTGTTTGA AGGAGGGGTG ACCGACTGT ATTACATCTC 420
 CAACACTCG AAAAGATCAT ACCCAACTC ATCCATCAC GTGACTGCG ACCAGTGTAC 480
 CATGGTCACC CAGCACGGGA AGCCCATGTT TACCAAGGTA TGTACAGAAG GCAGACTGAT 540

5 CTTGGAGTTC ACCTTTGATG ATCTCATGAG AATCAAAACA TGGCAGTTTA CCATTAGACA 600
 ATACCGAGAG TTAGTCCCGA GAAGCATCCT AGCCATGCAT GCACAAGATC CTCAGGTCTT 660
 GGATCAGCTG TCCAAAACA TCACCAGGAT GGGGCTAACA AACTTCACCC TCAACTACCT 720
 CAGGTGTGTG GTAAATTTGG AGCCAATGCA GGAAGTGTG TCGAGACATA AAACCTACAA 780
 CCTCAGTCCC CGAGACTGCC TGAAGACCTG CTTGTTTCAG AAGTGGCAGA GGAATGGTGG 840
 TCCGCGAGCA GAACCAACAA GGCACCAAC AACCACACGG AGAAAAGGA AAAATTCAC 900
 CAGCAGCACT TCCAACAGCA CGCTGGGAA CAATGCAAC AGCACTGGCA GCAAGAGAA 960
 GACCACAGCT TCCAACCTGA GTCTGTCCAG TCAGGTACCT GATGTGATGG TGGTAGGAGA 1020
 10 GCCAATCTGT ATGGGAGGTG AGTTTGGGGA CGAGGACGAA AGGCTAATCA CTAGATTAGA 1080
 AAACAGCAGT TATGATGCGG CCAACGGCAT GGAACGAGAG GAGGACTTCA ACAATTACCC 1140
 CGCGCTGGGG AACACAGGCC CGTGGAAACG TAAACCTCCC GCCACTCAAG AGACCAATC 1200
 AGAAAAACCC CCAACCCAGG CTTCCCAATA AGATGATCGG CACCAGAATC CACTGTCAAT 1260
 AGGCCCGTGG GTGATCATT CAATTGCAAA TCTTTACTTA CAGGAGAGGA AACAGAGAG 1320
 ATAAAAACTT TTCCATGCAA ATATCTATTT CTAAACACA ATGATCTGAT TTTCTTCTT 1380
 15 CTTTCTTTTT TTCTAATTGA GAGGATTATT CCCAGTAAGC TTCCATGACC CTTTCTTGGA 1440
 GGCCTTCACA GGTAAATACG ATACTGGCAC TGATTGTAAT TAAATGAGA GAAACTCTA 1500
 CGCATCTCTC TGGCAGGTT TTAACAAGT GTTGTGTTG AATTTCCTTT TTATGCATCA 1560
 AACGAAGGCC ATATTGTCCA TAAATGTCTA GTGCTCAGGA TCTCATTAAT ATGCGAACC 1620
 20 TAACACAGCA TGAATTTTTA ATATTGTAAA ATATTTCCTG CTTTCTGACT TGCATCTGAG 1680
 AGTTTCTGTG TCACTAATAA AAGAAAAAGA CAAAAAATC AGCTTTGAA AGTAATTTAA 1740
 ATGTACTCTA TTTTCTTTT CTTTATGTTT TCTTTCATTG GGCAACAGCT AAGAGGGCCC 1800
 AGCAAGGTAA TTTATGGTTG AGCTGATGTC AATTGGTTCT TGCTGTGAGT CGACTCAATT 1860
 TAGCCCAAGT GCTGAAACAA GAAATGTCTT TTTTTCATC AAGACACCA GGCAGATT 1920
 25 TTAAGTAAAG AAGACAAAT GAGCCCTTAA GAATTATGTC ATTTGTAAG TTGCTGTTGA 1980
 TCCAAATATT TTCAAGCCAT GTAATCCATT GGTTTGTGG GCAGTTTAAT AAACCTGAAC 2040
 CTTGTGTGTG TTTCTAATTG TACCTGAGTT GACCATCCTT TCTTTTATA GTATATTCT 2100
 TGATGATATG TTTGTAAAGC TCTCACCTGG TCTTTTATG GGGACTTTTC GTTTTGGGSC 2160
 AACTCCAGTG TATTTATGTG AAACTTTATA AGAGAATTAA TTTTCCATT TGCATATTAA 2220
 30 TATGTTCTCT CACACATGTA AAGGCACAGT GGTCCGCTGT GTTAAAAAC AGCTGTATTT 2280
 TATGTATGCT TACTGATAA GTGTGCCAAT AATAAAGT GTTAATGACC

Seq ID NO: 46 Protein sequence
 Protein Accession #: NP_001281

35
 1 11 21 31 41 51
 40 MSSTPHDPFYR SSFPGPFYRR HTPYMQPEY RIYEMNKRLQ SRTESDNLN WDAFATEFFE 60
 DDATLTLSFC LBDGPKRYTI GRLTLPYFS TVFEGGVDL YYILKHSKES YHNSITVDC 120
 DOCTMYTQHG KPMFTKVCTE GRLILEFTFD DLMRIKTWHP TIRQYRELVP RSILAMHAQD 180
 POVLDLQSLN ITRMLTNPFI LNYLRLCVIL EPMQELMSRH KTYNLSPRDC LKTCCLPQKQ 240
 RMVAPPAEPT RQPTTKRRKR KNSTSSSTNS SAGNNANSTG SKKKTTAANL SLSSQVFDVM 300
 45 VVGEPTLMGG EFGDEDERLI TRLENTQYDA ANGMDDEDF NNSPALGNNS PWNSKPPATQ 360
 ETKSENPPPQ ASQ

Seq ID NO: 47 Nucleotide sequence
 Nucleic Acid Accession #: NM_004126
 Coding sequence: 108..329 (underlined sequences correspond to start and stop codons)

50
 1 11 21 31 41 51
 55 GGCACGAGCT CGTGCCGGCC TTCAGTTGTT TCGGGACGCG CCGAGCTTCG CCGCTCTTCC 60
 AGCGGCTCCG CTGCCAGAGC TAGCCCGAGC CGGTTCTGG GGCAGAAATG CCTGCCCTTC 120
 ACATCGAAGA TTTGCCAGAG AAGGAAAAAC TGAAAATGGA AGTTGAGCAG CTTGCAAAAG 180
 AAGTGAAGTT GCAGAGACAA CAAGTGTCTA AATGTTCTGA AGAAATAAAG AACTATATTG 240
 AAGAACGTTT TGGAGAGGAT CCTCTAGTAA AGGGAATTC AGAAGACAAG AACCCCTTTA 300
 AAGAAAAAGG CAGCTGTGTT ATTTCAATAA TAACCTGGGA GAACTGCAT CCTAAGTGGA 360
 60 AGAAGTAGTT TGTTTAGTT TTCCAGATA AAACCAACAT GCTTTTAAAG GAAGGAAGAA 420
 TGAATTAATA AGGAGACTTT CTTAAGCACC ATATAGATAG GGTATGTAT AAAAGCATAT 480
 GTGCTACTCA TCTTTGCTCA CTATGCAGTC TTTTAAAGA GAGCAGAGAG TATCAGATGT 540
 ACAATTATGG AATAAAGAAC ATTACTTGAG CATGACACTT CTTTCAGTAT ATTGCTTGAT 600
 GCTTCAATAA AAGTTTGTGC TT

Seq ID NO: 48 Protein sequence
 Protein Accession #: NP_004117

70
 1 11 21 31 41 51
 MPALHIEDLP EKEKLMEVE QLRKEVKLQR QQVSKCSBEI KNYIBSRSGH DPLVKGIPED 60
 KNPFKEKGC VIS

Seq ID NO: 49 Nucleotide sequence
 Nucleic Acid Accession #: XM_051896
 Coding sequence: 139..2388 (underlined sequences correspond to start and stop codons)

	1	11	21	31	41	51	
5	GT	TTTAAAGA	CGCTAGAGTG	CCAAAGAAGA	CTTTGAAGTG	TGAAAACATT	TCCTGTAATT 60
	GAAACCAAAA	TGTCATTAT	AGATCCTTAC	CAGCACATTA	TAGTGGAGCA	CCAGTATTCC	120
	CACAAGTTTA	CGGTAGTGGT	GTTACGTGCC	ACCAAGTGA	CAAAGGGGGC	CTTTGGTGAC	180
	ATGCTTGATA	CTCCAGATCC	CTATGTGGAA	CTTTTATCT	CTACAACCCC	TGACAGCAGG	240
	AAGAGAACAA	GACATTTCAA	TAATGACATA	AACCCCTGCT	GGAATGAGAC	CTTTGAATTT	300
10	ATTTTGGATC	CTAATCAGGA	AAATGTTTGG	GAGATTACGT	TAATGGATGC	CAATTATGTC	360
	ATGGATGAAA	CTCTAGGGAC	AGCAACMTT	ACTGTATCTT	CTATGAAGST	GCGAGAAAAG	420
	AAAGAAGTTC	CTTTTATTTT	CAACCAAGTC	ACTGAAATGG	TTCTAGAAAT	GTCCTCTGAA	480
	TTTGTCTCAT	GCCCAGACCT	ACGATTTAGT	ATGGCTCTGT	GTGATCAGGA	GAAGACTTTC	540
	AGACAACAGA	GAAAAGAACA	CATAAGGGAG	AGCATGAAGA	AACCTTTGGG	TCCAAAGAAAT	600
15	AGTGAAGGAT	TGCATTCTGC	ACGTGATGTG	CCTGTGGTAG	CCATATTGGG	TTCAGGTGGG	660
	GGTTTCCGAG	CCATGCTGGG	ATTCTCTGGT	GTGATGAAGG	CATTATACGA	ATCAGGAATT	720
	CTGGATTGTG	CTACCTACGT	TGCTGGTCTT	TCTGGCTCCA	CCTGGTATAT	GTCACACCTTG	780
	TATTCTCACC	CTGATTTTCC	AGAGAAAAGG	CCAGAGGAGA	TTAATGAAGA	ACTAATGAAA	840
	AAATTTAGCC	ACAATCCCCT	TTTACTTCTC	ACACCAAGA	AAGTTAAAAG	ATATGTTGAG	900
20	TCCTTATGGA	AGAAGAAAAG	CTCTGGACAA	CCTGTCACTT	TTACTGATAT	CTTTGGGATG	960
	TTAATAGGAG	AAACACTAAT	TCATAATAGA	ATGAATACTA	CTCTGAGCAG	TTTGAAGGAA	1020
	AAAGTTAATA	CTGCACAATG	CCCTTTACCT	CTTTTCACTT	GTCTTCATGT	CAAACTGTAC	1080
	GTTCAGAGGC	TGATGTTTGC	AGATTTGGGT	GAATTTAGTC	CATACGAAAT	TGGCATGGCT	1140
	AAATATGGTA	CTTTTATGGC	TCCCGACTTA	TTTGAAGCA	AATTTTTTAT	GGGAACAGTC	1200
25	GTTAAGAAGT	ATGAAGAAA	CCCCTTGCA	TTCTTAATGG	GTCTCTGGGG	CAGTGCCTTT	1260
	TCCATATTGT	TCAACAGAGT	TTTGGGCGTT	TCTGGTTCAC	AAAGCAGAGG	CTCCACAATG	1320
	GAGGAAGAA	TAGAAAATAT	TACCAACAAG	CATATTGTGA	GTAATGATAG	CTCGGACAGT	1380
	GATGATGAAT	CACACGAACC	CAAAGGCACT	GAAATGAAG	ATGCTGGAAG	TGACTATCAA	1440
	AGTGATTAAT	AAGCAAGTTG	GATTCATCGT	ATGATAATGG	CCTTGGTGAG	TGATTGAGCT	1500
30	TTATTCATA	CCAGAGAAGG	ACGTGCTGGG	AAGGTACACA	ACTTCATGCT	GGGCTTGAAT	1560
	CTCAATACAT	CTTATCCACT	GTCTCCTTTG	AGTGACTTTG	CCACACAGGA	CTCCTTTGAT	1620
	GATGATGAAC	TGGATGCGAG	TGTAGCAGAT	CCTGATGAAT	TTGAGCGAAT	ATATGAGCCT	1680
	CTGGATGTCA	AAAGTAAKAA	GATTCATGTA	GTGGACAGTG	GGCTCACATT	TAACCTGCCG	1740
	TATCCCTTGA	TACTGAGACC	TCAGAGAGGG	GTGATCTCA	TAATCTCCTT	TGACTTTTCT	1800
35	GCAAGGCCAA	GTGACTCTAG	TCCCTCCGTT	AAGGAACCTC	TACTTGCAGA	AAAGTGGGCT	1860
	AAAATGAACA	AGCTCCCCTT	TCCAAGGATT	GATCCTTATG	TGTTTGATCG	GGAAGGGCTG	1920
	AAGGAGTGCT	ATGCTCTTAA	ACCCAAGAAT	CCTGATATGG	AGAAAGATTG	CCCAACCATC	1980
	ATCCACTTTG	TTCTGGCCAA	CATCAACTTC	AGAAAGTACA	GGGCTCCAGG	TGTTCCAGGG	2040
	GAACTGAGG	AAGAGAAGA	AATCGCTGAC	TTTGATATTT	TTGATGACCC	AGAATCACC	2100
40	TTTTCACCT	TCAATTTTCA	ATATCCAAAT	CAAGCATTCA	AAAGACTACA	TGATCTTATG	2160
	CACCTTCAATA	CTCTGAACAA	CATTGATGTG	ATAAAAGAG	CCATGGTTGA	AAGCATTGAA	2220
	TATAGAAGAC	AGAATCCATC	TCGTTGCTCT	GTCTCCCTTA	GTAATGTTGA	GGCAAGAAGA	2280
	TTTTTCAACA	AGGAGTTTCT	AAGTAAACCC	AAAGCATAGT	TCAATGACTG	GAAATGGCAG	2340
	CATTGCTGTA	TGCTGAGGCA	GTTTGCAATC	CCATGACAAC	TGGATTAA	AGTACAGTAC	2400
45	AGATAGTCGT	ACTGATCATG	AGAGACTGGC	TGATACTCAA	AGTTGCAAGT	ACTTAGCTGC	2460
	TAGAGATAAA	TACTATTATA	AGTTAGGTTG	ACAAATGATG	TTGATTATGT	AAGGATATAC	2520
	TTAGCTACAT	TTTCAGTCAG	TATGAACCTC	CTGATACAAA	TGTAGGGATA	TATACTGTAT	2580
	TTTTAAACAT	TTCTCACCAA	CTTCTTATG	TGTGTTCTTT	TTAAAAATTT	TTTTTCTTTT	2640
	AAAATATTTA	ACAGTTCAAT	CTCAATAAGA	CCTGCAATTA	TGTATGAATG	TTATTCACTG	2700
50	ACTAGATTTA	TTATACCAT	GAGACAAAC	TATTTTATT	TATATATGCA	TATATATACA	2760
	TACATGAAAT	AAATACATCA	ATATAAAAT				

Seq ID NO: 50 Protein sequence:

Protein Accession #: XP_051896

	1	11	21	31	41	51	
55	MS	FIDPYQHI	IVEHQYSHKF	TVVVLRAATKV	TKGAPGDMLD	TPDPVVELFI	STTPDSRKRT 60
	RHFNDINPV	WNETPEFILD	PNQRNVLBIT	LMADANYVMD	TLGTATPTVS	SMKVGEKKEV	120
60	PFIFNQVTEM	VLEMSLEVCS	CPDLRFSMAL	CDQKRTFRQQ	RKEHIRESMK	KLGLPKNSEG	180
	LHSARDVPVV	AILSGGGGFR	AMVGFSGVMK	ALYBSGILDC	ATYVAGLSGS	TWYMSTLYSH	240
	PDPPEKGPEE	INEELMKNV	HNPLLLTPQ	KVKRYVESLW	KKKSSGQPV	PTDIFGMLIG	300
	ETLIHNRMT	TLSSLKEKVN	TAQCPLPLFT	CLHWKPDVSE	LMFADWVBS	PYEIGMAKYG	360
	TFMAPDLFGS	KFFMGTVVKK	YEENPLHFLM	GVWGSAPFSL	FNRVLGVSGS	QSRGSTMREE	420
65	LENITTKYIV	SNDSDDSDDE	SHEPKGTENE	DAGSDYQSDN	QASWHRMIM	ALVSDSALFN	480
	TREGRAKGVH	NFMQLNLAT	SYPLSPLSDF	ATQDSFDDE	LDAVADPDE	FERIYEPLDV	540
	KSKKIHVVDS	GLTFNLPYPL	ILRPQGVDL	IISPDFSARP	SDSSPPFKEL	LLAEKNAMKN	600
	KLPPPKIDPY	VFDREGLKBC	YVFKPKNPDM	RKDCPTIIHF	VLANINFRKY	KAPGVPRRE	660
70	BEKRIADPDI	PDDPSFPST	FNQYVNPQAF	KRLHDLHMFN	TLANNIDVIR	AMVESIEYR	720
	QNPSCRSVSL	SNVEARRFFN	KRFLSKPKA				

Seq ID NO: 51 Nucleotide sequence:

Nucleic Acid Accession #: NM_006528

Coding sequence: 57..764 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
|      |      |      |      |      |
GCGGCCAGCG GCTTCTCGG AGSCCTTGCC CAGCGGGCCG CCGACCCCG TGCAACATGG 60
ACCCCGCTCG CCCCTGGGG CTGTCGATTC TGCTGCTTTT CTGACGGAG GCTGCACTGG 120
5 GCGATGCTGC TCAGGAGCCA ACAGGAATA ACAGCGAGAT CTGTCTCTCG CCCCTAGACT 180
ACGACCCCTG CCGGGCCCTA CTCTCCGTT ACTACTACGA CAGGTACAG CAGAGCTGCC 240
GCCAGTTCTT GTACGGGGGC TGGAGGGSCA ACGCCAACAA TTCTACACC TGGAGGGCTT 300
GCGACGATGC TTGCTGGAGG ATAGAAAAAG TTCCCAAAGT TTGCGGCTG CAAGTGAGTG 360
TGGACGACCA GTGTGAGGGG TCCACAGAAA AGTATTCTCT TAATCTAAGT TCCATGACAT 420
10 GTAAAAATT CTTTTCGGT GGGTGTCAAC GGAACCGGAT TGAGAACAGG TTTCCAGATG 480
AAGCTACTTG TATGGGCTTC TGCGCACCAA AGAAAAATCC ATCATTTTGC TACAGTCCAA 540
AAGATGAGGG ACTGTGCTCT GCCAATGTGA CTGCTATTAT TTTTAATCCA AGATACAGAA 600
CCTGTGATGC TTTCACCTAT ACTGGCTGTG GAGGGAATGA CAATAACTTT GTTAGCAGGG 660
AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGAAAAAGAA AAAGAAGATG CCAAGCTTTC 720
15 GCTTTGCCAG TAGAATCCGG AAAATTCCGA AGAAGCAATT TTAACATTC TTAATATGTC 780
ATCTTGTGTT TCTTTATGGC TTATTTCGCT TTATGGTGTG ATCTGAAGAA TAATATGACA 840
GCATGAGGAA ACAAATCATT GGTGATTAT TCACCAAGTT TTATTAATAC AAGTCACTTT 900
TTCAAAATAT TGGATTTTTT TATATATAAC TAGCTGCTAT TCAATATGTA GTCTACCAAT 960
TTTAATTATC GGTTCAACTG TTTGTGAGAG GAATTCCTGC AATGCATAAG ATATAAAGC 1020
20 AAATATGACT CACTCATTTT TTGGGGTCGT ATTCCTGATT TCAGAAGAGG ATCATAACTG 1080
AAACAACATA AGACAATATA ATCATGTGCT TTTAACATAT TTGAGAATAA AAAGACTAG 1140
CC

```

Seq ID NO: 52 Protein sequence
Protein Accession #: NP_006519

```

1      11      21      31      41      51
|      |      |      |      |      |
MDPARPLGLS ILLFLTEAA LGDAAQBEPTG NNABICLLPL DYGPCRAILL RYYDRYTQS 60
CRQFLYGGCE GNANNFYTWB ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFNLSSM 120
TCRKFPSGGC HRNRIENRFP DEATCMGFCA PKKIPSFCTYS PKDEGLCSAN VTRYVFNERY 180
RTCDAPFTYG CGGNDNMFVS REDCKRACAK ALKKKKKMPK LRFASIRKI RKKQF

```

35 Seq ID NO: 53 Nucleotide sequence
Nucleic Acid Accession #: AA478778
Coding sequence: no ORF found

```

40 1      11      21      31      41      51
|      |      |      |      |      |
TATTTTTGTA CGTAAATGA TTCTATTATG ACTGCCCTTG CATGTAGTAA TATGACAAAG 60
TGATCCCTCA TTATCAGCGT ACATATTGTT TTACTTTTCA TCTGTAAGTG TTTTATTGTT 120
45 ACTTTTAAAT AATGAATTTT TTTAAACAA TCTAGCCATC ATCAAGGTGC TATAAGAGTT 180
GTATAAAGAA TATTTTGGC ATTCTAGGC AAGTATCAGC CAATAAGTAT GTTAGTGATA 240
TCACGATTG TACCAACTAT TAACTATGTT AAATAAGTAT TCAATTTTCA GTGATCTCTG 300
GGAAAAAATAT ATGCTGCCTT GGTGCTAATA TTGTATGTAT TTAATATGAT ATCTGACTCA 360
GAAATATATA CACTTTTAAAT GAAAGGGAGG AAGGGAAGGA CAATTTCCAG TGCAAGAAAT 420
CACTTGGATG AAATAAGACC AGCTCTTTAC CCTTATTTTT GGATATGCCT TTTTGGGAAG 480
50 AGACTTAGAC TTTATCCTTA TTGTGTITAG TGTGTITAAAT ATTCGTTGCT TCAGCCCAAG 540
GTGCCTTGGT CTCTCCACAA TCAATGGAG GATCCCCCAA GCAGCTTCAT TACAGAGTGA 600
TATTTGGAAA GTGAGATCCT CTACCATTTT TGCCAAGATA CTCTAAAATG ACATCCAAGT 660
TTACCAAGTAG AAAGACACAG GATGCACAGA ATGGGCATGA CCTTCAGCTC ACAGACACAC 720
CTGGAGAAAT TCAGAACAGG GTTCTGAATC ATCAGGATG CCTTTTGATC GAAACATCG 780
55 GCTGGTGATG TGACTTCTCT TCAGGCCATG AGCCTAACAY CCTGCCGTTT TTTATGCCCG 840
CTGCAGTAAT GGACGTTTGT GTGAAGAAAT GAACTGTGGA GTACAAAATG CTTTGAGTCT 900
TTCCGATTTC TCAITTAATC ACTTTTGTG TACTTCTTTC CAAAATGGAA GTGCTGAAGC 960
CATGTCTTTT CTGCCCTCC AAGCTGATGA AGGGGAAGCCT TTGCCAATGG CCCATGGAAG 1020
60 ACATCTGGTT TGAGAAACCC TGCCCACTTC CAAAGACCAA AGAGATTAGG AAAAGCCTGG 1080
CAGTATTCTC CACTCCAAA CAAGCTCTAG AGTGCTCCAG GAAAAGTAT ATTCAATATA 1140
TGAATAGTGT TTATTTCCA TTATTAATGT GTTCTGAAAA TATATTATGA ATAAATACAT 1200
CACCACAGCC AAAAAAAAAA AAAAAAAAAA AAAA

```

65 Seq ID NO: 54 Nucleotide sequence
Nucleic Acid Accession #: NM_020663
Coding sequence: 1..645 (underlined sequences correspond to start and stop codons)

```

70 1      11      21      31      41      51
|      |      |      |      |      |
ATGAAGTGA AAGAGGGAAC TGACAGCAGC TGCGGCTGCA GGGGCAACGA CGAGAAGAAG 60
ATGTTGAAGT GTGTGGTGGT GGGGACCGGT GCCGTGGGGA AAACCTGCCT GCTGATGAGC 120
TACGCCAAGC ACGCCTTCCC AGAGGAATAC GTGCCCACTG TGTTTGACCA CTATGCAGTT 180
75 ACTGTGACTG TGGAGGCAA GCAACACTTG CTGCGACTGT ATGACACCGC GGGCAGGAG 240
GACTACACC AGCTGAGGCC ACTCTCTTAC CCCAACACGG ATGTGTTTTT GATCTGCTTC 300

```

TCTGTCTGTA ACCCTGCCTC TTACCACAAT GTCCAGGAGG AATGGGTCCC CGAGCTCAAG 360
 GACTGCATGC CTCACGTGCC TTATGTCTCT ATAGGGACCC AGATTGATCT CCGTGATGAC 420
 CCAAAAACTG TGGCCCGTTT GCTGTATATG AAAGAGAAAC CTCTCACTTA CGAGCATGGT 480
 GTGAAGCTCG CAAAGCGAT CGGAGCACAG TGCTACTTGG AATGTTTCTG TCTGACTCAG 540
 5 AAAGGTCTCA AAGCGGTTT TGATGAAGCA ATCTCACCA TTTTCCACCC CAAGAAAAAG 600
 AAGAAACGCT GTTCTGAGGG TCACAGCTGC TGTTCATTA TCTGA

Seq ID NO: 55 Protein sequence:
 Protein Accession #: NP_065714

1 11 21 31 41 51
 MNCKEGTSS CGCRGNDEKK MLKCVVVG DG AVGKTCLIMS YANDAPPEEY VPTVPDHYAV 60
 15 TVTVGGKQHL LGLYDTAGQE DYNQLRPLSY FNTDVFLICF SVVNPASYHN VQSENVPELK 120
 DCMHFHVTFL IGTQIDLRDD PKTLARLLYM KEKPLTYEHG VKLAIAIGA Q CYLBSCSALTQ 180
 KGLKAVFDEA ILTTPHFKKK KKRCSEGHSC CSII

Seq ID NO: 56 Nucleotide sequence:
 Nucleic Acid Accession #: fgenesh prediction
 Coding sequence: 1-546 (underlined sequences correspond to start and stop codons)

25 1 11 21 31 41 51
 ATGSCCTTGG GCAGCTCCGC CCCTGTGGCT TTGCAGGTA ATGCCCACTT CCTGTCTGCT 60
 TTCATGGCTG GCATTAAGTG TCTGTGGCTT TTCCAGGTAG TCCCTCTGGG GCTCCCGAG 120
 TTGGTCAAAA GGCTCCTGGG TGGAGCTCGA ACTGAAACTC GCTTTGTGCC CGCAGCCCTG 180
 30 CAGCTCGCCG GTGCCCTCGA CCTGCCCGCT GGTCTCTGTG CCTTTGAAGA GAGCACTTGC 240
 GGCTTTGACT CCGTGTGGC CTCTCTGCGG TGGATTTTAA ATGAGGAAGG CCAGCAACCT 300
 TTCTGTGCTC CAGGAGACAT GTCTGACTGG GACTACTGGG TTGGCTGGCG GAAGTTAATT 360
 CATCTCTCTC TGAGCACTCC AGGCTGAGAGC AGGCAGGTTA GGCTCCAGTT GTTCCAGCTT 420
 CAGTTTGTCA AAGGCCAGAA CTTGGACGTA CAGATGTACT GCAGGCTCCA GGGCAGTGAG 480
 35 AAACCCCTTG AAATGGTTC CATGGTTCCA TTCACCTTCA TGTACTGGAT CCACCATGGA 540
 AAGTAG

Seq ID NO: 57 Protein sequence:
 Protein Accession #: fgenesh prediction

1 11 21 31 41 51
 45 MALGSSAPVA LQGNAPFPAA FMAGIKLWL PQVVPGLLPE LVQRLLGGAR TETRFVPAAL 60
 QLQALDLPA GSAPPESTC GFDSVLASLP WILNEEGQQP FWSSGDMSDW DYMVGWRKLI 120
 HSPLSTPGWS RQVRLQLFQL QPVKGQNLDV TVYCRLLQSE KFPETGSMVP FTFMYIHHG 180
 K

Seq ID NO: 58 Nucleotide sequence:
 Nucleic Acid Accession #: XM_050478
 Coding sequence: 27..4508 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 55 CCGCGGGGCG CTGAGCCGAG CCGAGGATGG AGAACCGGCC TGGGTCTCTC CAGTAOGTCC 60
 CTGTGCAGCT GCAAGGGGGG GCACCTTGGG GCTTCACCTT TAAGGGGGGT CTGGAACACT 120
 GTGAGCGCGT CACAGTGTCT AAGATTGAAG ATGGAGGCAA GGCAGCTTTG TCCAGAAAGA 180
 TGAGGACTGG TGATGAGCTG GTGAATATCA ATGGCACTCC ATTATATGCG TCCCGCCNAG 240
 AGGCCCTCAT TCTCATCAA GGCTCCTTCC GGATTTCTCA GCTGATTTGC AGGAGGAGGA 300
 60 ACGCCCTCTG CAGTAGCCCG CACTCATGGC ATGTGGCCAA GCTGCTGGAG GGATGCGCTG 360
 AAGCAGCCAC CACCATGCAT TTCCCTTCTG AAGCCTTCAG CTGTCTCTGG CATCTGSGCT 420
 GCAACACAAG TGACGTGTGT GTGCAGTGGT GTCCACTCTC CCGGCATTGC AGCACCGAGA 480
 AAAGCAGTAC CATTTGGCAGC ATGGAGAGCC TGGAGCAACC AGGCCAAGCC ACCTATGAGA 540
 GCCATCTGTT GCCTATTGAC CAGAACATGT ACCCTAACCA CGGTGACTCA GCCTACAGCT 600
 CCTTCTCGGC CAGCTCAAAT GCTTCTGACT GTGCCCTTTC CCTCAGGCCA GAGGAGCCAG 660
 65 CCTCTACAGA CTGCATCATG CAAGGCCGAG GGCCAACTAA GGCCCCCAGT GGCCGGCCTA 720
 ATGTGGCTGA GACCTCAGGA GGTAGTCGGC GCACCAATGG GGGCCACCTG ACCCCAGCT 780
 CTCAGATGTC ATCCCGTCCA CAGGAGGGAT ACCAGTCAGG GCCCGCCAAA GCACTCAGGG 840
 GCCCACCAA ACCTCCAGTG AGGCGGACA GCCTTCAGGC CTCCAGAGCC CAACTCCTCA 900
 70 ATGAGAGACA GCGCAGGGCA TCTGAGCCTG TGGTCCCCTT GCCACAGAAG GAGAACTGA 960
 GCTTAGAGCC TGTGCTACCC GCAAGGAACC CTAATAGGTT CTGTTCCTCT AGTGGGCATG 1020
 ACCAAGTGAC AAGTGAGGGC CATCAGAACT GTGAGTTTCA TCAGCTCTCT GAATCCAGCC 1080
 AACAGGGCTG TGAGCATCTA CTGATGCAGG CCTCAACCAA AGCTGTTTGA TCCCAAAAAG 1140
 CCTGTGACAG AGCTTCCAGC GTGGATTCCA ACCCACTCAA TGAGGCTTCT GCAGAGCTAG 1200
 75 CTAAAGCTTC TTTTGGCAGA CCTCCACATC TCATAGGACC CACAGGGCAT GCCTATAGTG 1260
 CCCCCTGAACA GCTGTGGACA TCCCACCTGC AGCATGTGCA CCTTGATACC AGGGGCAGCA 1320
 AAGGATGGA GCTCCACCC GTACAGGATG GGCACCAATG GACTCTGTCC CCTTTGCACA 1380

	GCAGCCACAA	AGGGAAGAAA	AGTCCATGCC	CCCCTACAGG	AGGAACCCAT	GACCAGTCCA	1440
	GCAAGAGAAAG	AAAGACCAGA	CAAGTGGATG	ACAGGTCTTT	AGTTTTGGGA	CACCAGAGCC	1500
	AAAGCAGTCC	CCCACATGGA	GAGGCTGATG	GACACCCCTC	AGAAAAAGGT	TTCTTGGACC	1560
5	CAAAACAGAA	AAGCAGAGCA	GCCAGTGAAT	TGGCCAAACCA	GCAACCTCT	GCCTCTGGCT	1620
	CCCTTGTTCA	ACAAGCCACG	GACTGTTCTT	CAACCACTAA	AGCAGCTAGT	GGCAGAGAGG	1680
	CAGGTGAAGA	AGGGGACAGC	GAGCCCAAGG	AGTGCAGCG	GATGGGTGGT	AGGCGAAGTG	1740
	AGGGAGCCCG	GGGCGCTCG	ATCCAAAACC	GGCGGAAGAG	TGAGCGTTTT	GCTACCAATC	1800
	TGCGTAATGA	AATTAGAGG	AGGAAGGCC	AGCTCCAGAA	AAGCAAGGT	CCCTTGTCAC	1860
	AGCTGTGTGA	CACTAAGGAG	CCAGTGGAA	AGACCCAGGA	GCCTCCAGAA	AGTCTCTCAC	1920
10	TCAGTGCCTC	TAACACATCT	CTTCTATCTT	CATGTAAAAA	ACCTCCAGC	CCCAGAGACA	1980
	AGCTCTTCAA	CAAAAGCATG	ATGCTCAGAG	CTAGGTCTTC	CGAGTGCCTC	AGCCAGGCC	2040
	CTGAGAGCCA	TGAATCTAGG	ACAGGCTTAG	AGGGAAGGAT	AAGCCCTGGC	CAGAGGCTTG	2100
	GCCAGTCTCT	TTTGGGCTTG	AACACCTGAT	GGAAAGCACC	TGACCCATCC	TCCTCAGACC	2160
	CTGAGAAAGC	ACATGCTCAC	TGTGAGTCC	GTGGAGGTCA	TTGGAGATGG	TCTCCAGAGC	2220
15	ATAATTCACA	GCCACTTTGT	GCAGCAGCCA	TGGAAGGCC	TTCCAACCCA	GGTGACACCA	2280
	AGGAATTGAA	GGCTTCTACT	GCTCAAGCTG	GGGAGGATGC	CATCCTCTTG	CCTTTTGCAG	2340
	ACAGAAGAAA	GTTCTTTGAA	GAGAGTAGCA	AATCCTTATC	TACATCTCAT	TTGCCAGGTT	2400
	TAACCACTCA	TAGCAACAG	ACTTTTACCC	AGAGACCAAA	ACCTATAGAC	CAAAACTTCC	2460
	AGCCAACTAG	CTCCAGCTGT	AGGGAATTGA	GGCGCAATCC	CACTGACCAA	TCATATCATT	2520
20	CCGAGAGCCA	ACCATATCAT	GCCACAGACC	AATCATATCA	TTCCATGTCA	CCCTTTCAGT	2580
	CAGAACTCC	CACCTTACTCA	GAATGTTTTG	CAAGCAAGG	TCTAGAAAAA	TCCATGTGTT	2640
	GTAAGCCACT	ACACTGTGTT	GATTTTGATT	ACCACAGGAC	CTGCTCTTAC	TCTTCAGTGT	2700
	TCAAGGAGC	TCTAGTCCAT	GATCCTTGCA	TTTATTTGTC	TGGGAAATTC	TGCCCTGCCT	2760
25	TGCTAAGAGG	AAATATGATG	CCAAATGTGT	ACAACTGCG	GTGCCACCA	CACCAATGCA	2820
	TTGCTGTGTC	AGTTTGCTAT	CATAATCTCT	AGCAGAGTGC	CCTGAGGAGC	AGCAGCTTGG	2880
	CACCTGGCAA	CACTTGGA	CCCAGGAAGC	TGACAGTGCA	GGAAATTTCT	GGGGACAAAT	2940
	GGAAATCCAAT	AACAGGAAC	AGGAAGACCA	GCCAGTCAGG	GAGGGAATGT	GCTCATTTCCA	3000
	AGACTAGCTT	TTCAATGGCA	ACCCCTTTCC	ATCCTTGCTC	TGAGAACCCA	GCACTGGAAT	3060
	TGTCAGAGCTA	CGAGCAATTT	TCTTCTCTTG	ACCTCTCTTG	AGACTTCAAA	CATGCTTTGA	3120
30	AAAAATCAGA	GGAACTTCA	GTTTATGAGG	AGGGGAGCTC	CCTTGCTTCC	ATGCCCAACC	3180
	CACCTGCGAG	CCGTGCTTTC	TCAGAGAGTC	ACATCAGCTT	GGCGCCCCAA	AGCACCCGGG	3240
	CCTTGGGGCA	GCATAGGAGG	GAGCTCTTTA	GCAAAAGTGA	TGAGACCCAG	TCGGATCTTC	3300
	TCGGAGCCAG	GAGAAGGCC	TTTCTCTCTC	CTCGCCCTCC	TTCTCCCAAC	TGGGAGAGAT	3360
35	ACAGGCTCTT	TCGTGAGGCC	CAGCAGCAGA	AGCAGCAACA	GCAGCAGCAG	AAGCAACAGG	3420
	AGGAGGAGGA	GGAGGAGGAA	GAAAGAGAAG	GAAAGAGAAG	GAAAGAGGAG	GAGGAGGAGG	3480
	CAGAGGAGGA	GGAGGAGGAG	CTGCCACCCC	AGTATTTTCA	TTCAAGAAAC	TCTGTTTCTT	3540
	GTGCTGCTTA	TCCTGAGGAG	GTCTTAGAGC	AGCCACAACC	CCTCAGCTTT	GGCCACCTGG	3600
	AGGGCTCGAG	CAGGGGTTCA	CAAAAGTTC	CAGCAGAGCA	AGAACTCTTT	GCACTCCATT	3660
40	CCAGTAGTTT	CTTGCTTCCA	ATAAGGGGTC	ACTTGGGATC	TCAACTGTAG	CAGGCTCAGC	3720
	CCCTTGCTTA	CTATGCGATT	GCTGCGCTTT	GGAGGACATC	GGGACAGGAA	GCCACTGAAT	3780
	CCGCCAAACA	AGAGTTTCAG	CACCTTTTCG	CTCCTTCAGG	GGCCCCAGGA	ATCCCTACCT	3840
	CTTACTCAGC	TTATTACAAT	ATTCTGTGAG	CCAGGCGAGA	GCTGTGAAC	AAACTGAAAG	3900
	ACCAACTGTA	GATGGCAGAG	ATTGCGCTAG	GAGAGGAGGA	AGTTGACCAT	GAACTGGCTC	3960
45	AAAAAAGAT	ACAGCTTATC	GAAAGCATCA	GCAGAAAACT	TTCTGTCTTG	CGGGAGGCC	4020
	AGCGAGGCTT	GCTAGAGGAC	ATCAATGCCA	ATTCTGCCCT	TGGGGAGGAG	GTGGAGGCCA	4080
	ACTTTAAAGC	CGTCTGCAAA	TCCAATGAAT	TTGAAAGATA	CCACTTGTTT	GTGGGGACCC	4140
	TGGAACAAGT	GGTCAACCTG	TTGCTGTGAC	TCTCTGGAGC	ACTGGCCCGG	GTGGAGAAATG	4200
	CTCTGAACAG	CATGATTCCA	GAGGCCAAC	AGGAGAGGTT	GATACTGATA	GAGAAGAAGC	4260
50	AGCAGCTGAC	GGGCGAGTTG	GCAGATGCCA	AGGAGCTGAA	GGAGCAGCTG	GACCGCCGGG	4320
	AGAAGTTGGT	GTTTGGCATG	GTCTCCCGCT	ACCTGCCCTCA	GGAGCAGCTC	CAAGATTACC	4380
	AGCAGCTTGT	CAAGATGAAA	TCGTCTCTCA	TCATTGAACA	GGGAGAGCTG	GAGGAGAGGA	4440
	TCAGCTCGG	GGAGAGGCAA	CTCAAAATGT	TCAGGGAGAG	TCTACTCTCT	GGGCCAGCA	4500
	ATTTCTAAT	CTACAGCAC	CTGCCCACAG	CATCCCTGCC	CAGCCATGTG	GGAAAGTCTT	4560
55	TCAATCTTCT	TTGTTAGCAG	TTTCTCAGCA	AGTAGATAGC	AATTAGCAGT	TTGTTCCAGC	4620
	CCTCTACCTT	GGATGTCTCT	CACCTACCTT	TCCCTAGCAG	TGGTCTTAAC	CAGCTAGGAG	4680
	ACCCTGGGGA	AGCCACAAGC	TTCTACCCAA	GGGAGCTGCA	GCAAGGTGTG	ATCTTAGAAC	4740
	CACACTCTCT	TTCCACAGT	TGCCAAGGGC	AAGTACTTGC	TGCACAGAGA	ACCAAGGAAG	4800
	TGCTTCTATT	CTGCTTTGTA	CTAGGACACC	AAAGACATCA	AGTACTCATC	ACCCACCCAT	4860
60	ATCATCAACA	GCCTCTAAAG	GCTCAGAGGG	AATCTGCCTT	GCAGCTCTAC	TCTGCCCCAG	4920
	GGCTTGTGGC	CAGCCATTTC	TCACAGAGAG	CTGGCTGCCT	TGAGGGCAAT	CACCTGACAC	4980
	CAGTTTCAGG	GCCTACCCCA	AGCTTTGAG	GGGAAAGCAC	AGAGGGAGGA	ATTACACTGA	5040
	AAAAAATGCA	AGCAAAAGTT	GAGTACCCCC	AGGTGCCCTT	TAGGAAGGAA	CCAGGTTTAA	5100
	ATAGGCTCTA	CCCTTACCTT	TCCAGCAGC	AAGTTACAGG	GAAAGGGCCT	ACTCTTAGCC	5160
65	CTGGCTAGTG	TGACCTCTTT	CCTGTCTTAA	GACTTTGGTC	CTACCACTTC	TTGTTTATC	5220
	TTTCTTTTAC	ATTGCTGGGG	GTTCACCGAG	GTGCTTACCC	CAGGGCTTCA	CCATATGGGC	5280
	CATTAAATAG	TCTATAAAAA	CTGACTTCTA	GATGTAGGTT	TCATTATTGG	GGGAGGGGGT	5340
	TCTTATTGTT	ATATTITAAA	TGCGCTTTTG	ATTTTATTIA	TTTTTATGTT	TTGATTATTT	5400
	TTTTCTTTTT	TAACTAATAA	GGGAGGAAGA	GGGAAGTTGG	AGAGGGAAAA	GTTAGCCCGG	5460
70	AAGGAAAGCA	TTTTCTGAG	ATCAGCCTGA	ATCCACCGTG	GCTAGGCATA	TTCTTGCTCT	5520
	TCTCGTGTGG	CTCACAACTA	CTGCTGTGGA	TGAATTTAGG	AAAGTTGCAG	GATACAGAGT	5580
	TAAACACAAA	GATCAATGAA	ACAAATCCGA	AATGTTATTA	AGAAACAGT	TCGGCCGGGG	5640
	CATGTTGGCT	CAGCCCTGAA	ATCCAGCAC	TTTGGGAGGC	CGAGGAGAGT	GGATCAGGAG	5700
	GTCAAGAGAT	CAGAGCAATC	CTGGCTAACA	CGGTGAAACC	CTATCTCTAC	TAAAAATACA	5760
75	AAAAATAGC	CAGGTGTGGT	GGCAGCAGCC	AGTAGTCCCA	GCTACTCGGG	AGGCTGAGGC	5820
	AGGAGAAATG	CTTGAACCTG	GAAGGCAGAG	ATTGCAAGTG	GCTGAGACCA	CACCACTGCA	5880
	CTCATCTCTG	GGCAACAGAG	TGAGACTTTG	TCTCAAAAAG	AAAGAAAGAA	AGAAAGAAAG	5940

AAAGAAAGAA AGAAAGAAA GAAAGAAAGA AAGAAAGAAA ACAGTTCCAT TTACATAGC 6000
ATC

Seq ID NO: 59 Protein sequence:
Protein Accession #: XP_050478

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75

1	11	21	31	41	51	
MENRPGSFQY	VPVQLQGGAP	WGPTLKGGLE	HCEPLTVSKI	EDGGKAALSQ	KMRTGDELVN	60
INGTPLYGSR	QEALILIKGS	FRILKLIVRR	RNAPVSRPHS	WHVAKLLEGC	PEAATTHFP	120
SEAFSLSWHS	GCNTSDVCVQ	WCPLSRHCST	EKSSSIGSMB	SLBQPGQATY	ESHLLPIDQN	180
MYPNQRDSAY	SSPSSASSNAS	DCALSLRPEE	PASTDCIMQG	PGPTKAPSGR	PNVARTSGGS	240
RRTNGGHLTP	SSQMSRRPQS	GYQSGPAKAV	RGPPQPFVRR	DSLQASRAQL	LNGEQRRASE	300
PVVPLPKKSK	LSLEPVLPAR	NPNRPFCLSG	HDQVTSBGHQ	NCFBQPPFES	SQQQSEHLLM	360
QASTKVGSP	KACDRASSVD	SNPLNEASAE	LAKASFGRPP	HLIGPTGHRH	SAPBQLLASH	420
LQHVHLDTRG	SKGMELPFVQ	DGHQWTLSP	HSHKGGKSP	CPPTGGTHDQ	SSKRRKTRQV	480
DRSLVLVGHQ	SQSSPPHGEA	DGHPSKGGFL	DPNRTSRAAS	ELANQQPSAS	GLVQQATDC	540
SSTTKAASGT	EAGEBGDSEP	KRCRMRGGR	SGOTRGRSIO	NRRKSERFAT	MLRNBQRRK	600
AQLQKSKGFL	SQCDTKKEPV	EETQEPPEEP	PLTASNTSLL	SSCKKPPSPR	DKLFNKSHML	660
RARSSECLSQ	APESHESRTG	LEGRISPGQR	PGSSSLGLNT	WMAKDPSSS	DPEKAHRCG	720
VRGGHWRWSP	EHNSQPLVAA	AMGSPSNPGD	NKELKASTAQ	AGEDAILLFP	ADRRKPFPEES	780
SKSLSTSHLP	GLTTHSNKTP	TQRPKPIDQN	PQPMSSSCRE	LRRHFMDSY	HSADQFYHAT	840
DQSYHMSMPL	QSBTPPYSEC	PASKGLENSM	CKPLHCGDF	DYHRTCSYSC	SVQGLVHDP	900
CYKSGSICFP	ALLKRNMMPN	CYNCRCHHQ	CIRCSVCYHN	PQSHLEDS	LAPGNTWKPR	960
KLTQVEPPGD	KWNPITGNRK	TSQSGREMAH	SKTSFWSWATP	FHPCLNPAL	DLSSYRAISS	1020
LDLLGDFKHA	LKSEBTSVY	EGSSSLASMP	HPLRSRAFSE	SHISLAPQST	RAWGQHRREL	1080
FSKGDFTQSD	LLGARKAFPP	PPRPPPNWME	KYRLPRAAQ	QKQQQQQKQ	QSEEEEEEE	1140
EEEEEEEEEE	EEEEEEEEEL	PQYPSSETSG	SCALNPEEVL	BQPOLSPGH	LESGRQSGS	1200
VPABQESFAL	HSSDFLPPIR	GHLQSQPEQA	QPPCYIGIG	LWRTSGQZAT	ESAKQEPQHF	1260
SPPSGAGPIP	TSYSAYYNIS	VAKASLNLKL	KDQPEMASIG	LGBEEVDHEL	AQKKIQLIES	1320
ISRLSLVLE	AQRLLEDIN	ANSALGHEVE	ANLKAUCKSN	EPEKYHLPVG	DLDKVNNLL	1380
SLSRLARVE	NALNSIDSEA	NQELVLIEK	KQLTGQLAD	AKELKEHVDR	REKLVPQMVH	1440
RYLPQDQLQD	YQHPVMMKSA	LIIQRELES	KIKLGESEKL	CLRESLLQGF	SNF	

Seq ID NO: 60 Nucleotide sequence:
Nucleic Acid Accession #: NM_014705
Coding sequence: 192..2489 (underlined sequences correspond to start and stop codons)

40
45
50
55
60
65
70
75

1	11	21	31	41	51	
GGGAGAGCT	AGGAAAAAT	GTCTTTGAGC	TGTGAGATGC	TGTATATATT	TGAAAAATATG	60
ATTATATGCA	TGTGTTTGTA	TTTTATGACT	TGGATAATCT	GAAAAATCAAT	TTGCTTTGTC	120
AATGCTTCCT	GGATTAGAAAT	TCCACTATT	GGTCCCTATC	CTAGTCTACT	AAAGAAAAAT	180
GAGCGGGAAA	CATGGCGGGA	AAGTGGCGTT	TCATTAAATTG	CTACTGTAAC	TGCTCTAATG	240
GAGAGGTGT	TAGATTACAG	AACCTCTATA	AGACTGAAC	GAACAAGGAG	GAGATGTATA	300
TACGCTACAT	TCACAACTC	TATGATCTGC	ATCTCAAAGC	ACAGAACTTT	ACAGAACTG	360
CATATACCTC	CCTCTTATAT	GACGAGCTAC	TGGAAATGGTC	TGATCGGCC	CTCAGGAGT	420
TCCTGACCTA	CCCCATGCCA	ACAGAAATGGC	AGCGCAAAGA	GCACCTGCAC	CTCACCATCA	480
TCCCAAGACT	TGACAGAGGC	AAATGTTGGG	AGAATGGCAT	TATCTTGTGC	CGGAAGATTG	540
CAGAGCAGTA	TGAGAGTTAT	TATGACTACA	GAAACCTGAG	CAGATGCGG	ATGATGGAAG	600
CCTCTTTGTA	TGACAAAATT	ATGGACCAGC	AACGTCTTGA	ACCAGAGTTC	TTCAGAGTTG	660
GATTTTATGG	AAAAAAATTT	CCATTTTCT	TAAGAAATAA	GGAGTTTGTG	TGTGAGGGGC	720
ATGACTACGA	GAGGCTGGAA	GCCTTCCAAC	AGAGAATGCT	GAACGAGTTC	CCCCATGCCA	780
TGCGCATGCA	GCAAGCCAAC	CAGCCCGATG	AGACCATCTT	CCAGGCAGAA	GCTCAGTATT	840
TGCGATATA	TGCTGTGACT	CCCATTCAG	AGAGCCAGGA	GGTCTGCG	AGAGAGGGTG	900
TTCCGACAAA	CATCAAAAGC	TTCTATAAAG	TGAATCACAT	CTGGAATTC	CGCTATGACC	960
GACCATTTCA	CAAGGCCACA	AAAGATAAAG	AGAATGAATT	CAAGAGTCTC	TGGGTGGAGA	1020
GACGTCATT	ATACTTGCTG	CAGAGTTGCG	CTGGCATCTC	TGCTGCTTT	GAAATGGA	1080
AGCGTGAAGT	GGTAGAAATG	AGTCTCTGCG	AAATGCAAT	TGAAGTGCTA	GAAATAAGA	1140
ATCAGCAGCT	GAAGACTCTG	ATTAGTCAAT	GTCAGACAAG	ACAGATGCG	AAATATTATC	1200
CCCTGACTAT	GTGCTGAAT	GGAGTTATAG	ATGCTGCAGT	TAATGTTGCG	GTTTCCAGGT	1260
ATCAAGAGGC	ATTCTTTGTC	AAAGAAATATA	TCTTAAGTCA	CCCTGAAGAT	GGGAGAGAAA	1320
TTGCACGATT	AAGAGAGCTG	ATGCTTGAGC	AGGCACAGAT	TCTGGAATTT	GGTTTGGCCG	1380
TGCTATGAGAA	GTTTGTACCT	CAAGATATGA	GACCCCTTCA	CAAAAAGCTG	GTTGACCAAT	1440
CTCTTGTGAT	GAGTTCAGAGC	TTAGGGATAC	AGGAGTTCTC	TGCTTGTATG	CAAGCCAGTC	1500
TGCTTCAATT	TCCTAATGGA	AGCCCTCGTG	TGTGTAGAAA	CTCAGCACCT	GCTTCTGTGA	1560
GCCTCAGATG	TACCAAGGTA	ATTCCTAGAC	GCAGCCCGTG	AGTTTACCCA	GCTGTCAACC	1620
GATATTCTTC	CTCCTCACTG	CTCTCACAG	CTTCTGCTGA	AGTAAGCAAT	ATTACAGGGC	1680
AATCAGAAAG	CTCTGATGAA	GTCTTTAACA	TGCAGCCAG	TCCATCTACC	TCAAGCTTGA	1740
GTCTTACTCA	CTCGGCTTCA	CCTAATGTGA	CAAGTTCTGC	TCCATCGAGT	GCCAGAGCTT	1800
CTCCTTTGTT	GTCTGACAAA	CACAAACATT	CCCGAGAAAA	CTCTTGCCTG	TCACCAAGAG	1860
AGAGACCATG	CAGTGCCATC	TATCCAACAC	CTGTGGAGCC	TTGCGCAGAG	ATGCTGTTTA	1920
ATCATATTGG	AGACGGGGCC	TTGCCACGCA	GTGACCCAAA	TCTCTTGCA	CCTGAAAAAG	1980
CTTCCACGAC	AAACACACG	ACATCAGTAT	CCCTCTCGCC	TGCGGGCGGA	TCTCATTGTA	2040

5 AGGGCTCTGT CGAGTCTTTC ACCCCCTCTC CAGTGGAGTA CCACTCGCCA GGACTCATCT 2100
 CCAACTCCCC TGCTTGTGCG GGCAGCTACA CGAGTGGGAT TTCTTCTCTC AGCCGGTGCA 2160
 GCACGTGCGA AACCTCAGGC TTGAAAATC AGGTGAATGA ACAGTCGGCC CCCTGCGCGG 2220
 TGCCAGTGCC GGTGCCCGTG CGGAGCTACG GCGGGGAGGA GCCAGTGGCG AAGGAGAGCA 2280
 AGACTCCGCC CCGTACAGC GTCTACGAGC GGACTCTGCG GCGCCCGCTC CCGTACCTC 2340
 ACAGCTCTCT CATCCCGTTC ACGTCGGAGC CGCCCGGCTC GCGCCCGAAG CCTCTGGCAG 2400
 CGGATCCAGC CCACTCGGAG AATGGGGCCC GGAGGACTGA CCGCGGCCCG CGGCCCAGGC 2460
 CCCTGCCCGG CAAGTCTCT CAGTTATAG TCACCTTCTT ATGTACCTGC GATGCAATTCT 2520
 10 TTGCCCGTTT ACAAATAAG AAGTATGATG AGAAGACATT TAGTGTAGGC ACTTTAATAA 2580
 TCCTACGAGC TCCTTCGATG AATGGAATTA AAAGTGCCTT ATTAAATATC ATGTGACACA 2640
 ATATTAAAGT TTGCTGATCT AAAACGCCAG ATGTAAATG AAGTATGGCT GAATTTCATT 2700
 AAAACGTTTC TCATTGGGAA GTGGTAAATA GTGATAAAGA CTCCTTTTGT ACCTTTTAT 2760
 GTTCACTTTT TTTTATATAG TTTAATCTTA AAACCAATAC GATATTGTCA AACGATACAA 2820
 TGTTGTACAA TGTTGTATCG TTTTACTGTA ATACTTGATA CTGTGAGAAA GCTTATTAA 2880
 15 TCCTGACACA TCCTAACACA GTGGTCTTTA TTTTGAAGA CTTCGTGAAA TAAGGCAAGG 2940
 TTTATCAGTG CAGATCATCA GAATTAAGT TCAAGCAGGC GAGCAAGACA GTTACTTAA 3000
 GGGGTGTCAA AGCTTGGGAC TGGAAATTGT TTTGTCTTG AAACAAAATA CTCTTTTAA 3060
 GTTCACTTTT CTTTGTGACT GCTGTCTACA TTCTGAAAT TCTATTTTGT GAATTGGTAG 3120
 20 TAAATCCCTT TACTACCTCG ACACCGTGGT ATCTACTGTA TTCTTTTCA AGGTGCAATT 3180
 TGCTTCAGAG TTCCAATCAG CTAGATTAA 3240
 ATTTTCGCTC TCCTTCTTCT ATAGTTTCTT ATATAAAAT TGTCATTGAA CAAGAGCAAA 3300
 TGCTGAAGTA TTAATGAGGC ACAAATGACT GTGCCCCATT AGCAAGAATT CAGGAATCAA 3360
 TACAGACAGT ATTAAATTA TAGCTTAAGT GAAGAAAAA AAAAAGTAA TGAATATGTA 3420
 25 TTAGCAGCAT TAAATGGCAA AAGGACTTAT AAAAGGCAAG GGCATTAACT TTCACTCTG 3480
 CACAAATAAA AAAATCTCT ACGACTCTCC ACTTTTACCA GTGGAGTTTG TCTTAGCTGA 3540
 CCTGTGCTCT TTCTCTGAA GGAGGATTGC TGTAGACTTC TCTAGCTTGA ATATTGCAAC 3600
 ATAGCATCTT AGGTCTAGAT AGGGATGCTA ATGCCAGTTG TAGAAGTGTG AAAAAGCAC 3660
 CTGTATGTA GTAATGATT TTATATCTTT GTTTTCTT TTACTGACTG TTTTAAACAC 3720
 TCAATGACA ATAGATATGA ACTGATTTT AAATCACT GTTAAATATT TTCCCTCTTT 3780
 30 TGTGCGAAG CTCAATTAG TTTAACCATG TTTGTTTGT TGGTAGCTTA CTGGAAGGC 3840
 ACTGACCACT TTTTATATT CTCCTAATGA AACCATTCAG CAGGTATATG CTGTGAGGC 3900
 TGGTATAGA GGTTTTCTAT AATAAATGTT CAGTATTTT TGTATATAAC TGGTAAATTT 3960
 TAATAGAGTA TACCATATG TGTAAAAAAGT AAGCAAAACA GTTGTATGAT 4020
 35 CAGTATGATT GTTATAATTA TGCCAAATAC TTTACGTATG GAAAGAAGAT ATTTGTACAT 4080
 TATTCTGTTT GCTTTAAAGA AGTAAACCT TCCATTAA TAAGATGACA TGCATAAGAT 4140
 ATATGTTAAA TATTTATGTT TAGTGAAAGT GTTCATAATT GAGAAAAGGA ACATATGCAT 4200
 TATCTCTTG TATCTGCAA GTTTGCAAT CAGAAATTTT TTGAAGTACG TTTTGTCTTT 4260
 GATAACACTT CGTGTGTA ACCACATTC TATATATATA CATATATATG TGAAGCTCCA 4320
 40 TTTTCTGTTT GCTTTAAAGA AGTAAACCT TCCATTAA TAAGATGACA TGCATAAGAT 4380
 AACAAAGCTT CCGTATTTC CTTTCTCTGT GTAATTAAAT AGATTGTGT ACTAGTGCTT 4440
 GGGCAGACTA TAAATCAGTG TTATTTGCTC TTGGAGCCAT TTTTAAAAA AAATTTTGGC 4500
 AGTGACAGT TGAATTTATC TTGAATTAT CATGTGTGTG TATTCTGAA GCAGCTACAT 4560
 AGCAGAACAT TTTAAGAGAT TCTGTAGGCC CACATGTTCA TGTGTGTTGC TGCTGAATGG 4620
 45 TAAATATTA AATAAATAC CAGATTAATC TT

Seq ID NO: 61 Protein sequence
 Protein Accession #: NP_055520

50 1 11 21 31 41 51
 MAGKMRPINC YCNSSNGEVV RLQNFYKTEL NKEEMYIRYI HKLYDLHLKA QNFTEAAAYTL 60
 LLYDELLEWS DRPLREPLTY PMQTEWQRKE HLHLTIQNF DRGKOWENGI ILCKRIASQY 120
 ESYVDYRNLK KMRMMEASLY DKIMDQQRLE PEFFRVQFYG KKPFPFLRNK EFCVCRGHDYE 180
 55 RLEAFQQRML NEFFHAIAMQ HANQPDFTIF QAEAYLQIY AVTFIPBSQE VLQREGVFDN 240
 IKSFYKVNHI WKFRYDRPFH KGTIDKENEK KSLWVERTSL YLVQSLPGIS RWFVEBKREV 300
 VEMSPLENAI EVLENKQQL KTLISQCQTR QMKNINPLTM CLNGVIDAAV NGGVSRYQEA 360
 PPFVEYILSH PEDGEKIARL RELMLEQAQI LEFGLAVHEK FVPQDMRPLH KKLVDQFFVM 420
 KSSLIGQEFBS ACMQASPVHF PNGSPRVCRN SAPASVSPDG TRVPIRRSPL SYPVNRYS 480
 60 SSSLSSQASAE VSNITGQSES SDEVFNMQPS PSTSSLSTH SASPNVTSBA PSSARAGPLL 540
 SDKHKHSREN SCLSPRRPC SAIYPTPVEP SQRMLEFNHIG DGALPRSDPN LSAPKASPA 600
 RHTTSVSPSP AGRSPLKGSV QSFTPSFVEY HSPGLISNSP VLSGSYSSTG SLSRCSTSE 660
 TSGFENQVNE QSAPLPVFPV VFPVSYGGBE FVRKESKTPP PYSVYERTLR RPVPLPHSL 720
 IPVTSBPAL PPKPLAARSS HLENGARRTD PGPFRPLPLR KVSQ

65 Seq ID NO: 62 Nucleotide sequence
 Nucleic Acid Accession #: fgenesh prediction
 Coding sequence: 1..2561 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 ATGACCGAG GCCAGGGTAA GAGGGGCGGC GACGCGCGCA CTGTGTTGCG GCGCGGGCGG 60
 GAAAGGAGGA CTGACGATC TGAAGCCGGA GAGGAGGAGC GAGAGAGGCG GCGCGTGGGG 120
 CGGGGCGTGA GGAACGCTCG GAGGGGACTG GAGAGCGCGG CGCTTATGCA AAGGTGCTTT 180
 75 CGGCTGCGCG GACAAACCGC CAGCAACCAG GTACAGCTCT CAGAGGTTC CAGAGGAGAG 240
 CTCAGGTGCC CTGAATCTCC CAGTGTGGCA GAGAAAGTGA AACTTGTGCA CCGATGCTCG 300

5 GAACTGCTGG AGCAGCTGCT CCCAGAGCTC ACCGGGCTGC TCAGCCTCCT GGACCACGAG 360
 TACCTCAGCG ATACCAACCTT GGAAGAGAG ATGGCCGTGG CCTCATCCT GCAGAGCCTG 420
 CAGCCCCCTC CAGCAAGGA GTCTCTCTAC CTGTATGTGA ACACAGCAGA CCTCCACTCG 480
 GGCCCCAGCT TCGTGGAAATC CCTCTTTGAA GAATTTGACT GTGACCTGAG TGACCTTCGG 540
 GACATGCCAG AGGATGATGG GGAGCCGAGC AAAGGAGCCA GGCCTGAGCT AGCCAGAGGC 600
 CCAGCCCTGA GAAAGCGGCG CGACCTGCCT CCACCGCTCC CCAACAAGCC TCCCCCTGAG 660
 GACTACTATG AAGAGGCCCT TCCTCTGGGA CCGGCAAGT CGCTGAGTA CATCAGCTCC 720
 CACCAATGGCT GCAGCCCTC ACACCTGATT GTGATGGCT ACTATGAGGA CGCAGACAGC 780
 AGTACCTCG CAACAGGGT GAACGCGAG CTTAAGAGCT CCTATAATGA CTCTGACGCA 840
 10 ATGACGAGT CTTATGAGTC CTACGATGAA GAGGAGGAG AAGGGAAGAG CCGCAGGCC 900
 CGACACCAAT GGCCTCAGA GGAGGCTCC ATGCACCTGG TAGGGGAATG CAGGATATGT 960
 GCCTTCCTGC TCGGAAAAA GCGTTTCGGG CAGTGGGCCA AGCAGCTGAC GGTTCATCAG 1020
 GAGGACCAGC TCTGTGTGA CAAAAGCTCC AAGGATCGGC AGCCACATCT GAGGTTGGCA 1080
 CTGGATACCT GCAGCATCAT CTACCTGCCC AAGGACAGCC GGCACAAGAG GCACGAGCTG 1140
 15 CTTTTCACCC AGGGGCTAC CGAGGCTTTG GTGCTGGCAC TGCAGAGCCG AGAGCAGGCC 1200
 GAGGATGGC TGAAGTTCAT CCGAAGAGTG AGCAAGCCAG TTGGGGGAGC TGAGGGAGTG 1260
 GAGGTCCGCA GATCCCACT CCTCTGTGC AAGTTGGACC TGGACAAGAG GCTGTCCCAA 1320
 GAGAAAGCAG CTTGAGATTC TGACAGCGTG GGTGTGGGTG ACACTGTTC TACCTTGGC 1380
 CGCGGAGGAG CTTGTGATCA CGCAAGGGG AAGAAGAGCA GCTTGCAGA AGTGAAGGCC 1440
 20 TCAATGAGCA GGGCTGCGGG CCGCAAGATC ACCCTATCA TTGGCTTCTC CAAGAAGAG 1500
 ACACTGSCCG ATGACCTGCA GACCTCTCC ACCGAGGAG AGGTTCCCTG CTGTGGCTAC 1560
 TGAACCTGCT TGTGAACCA GGCCTGGAAG GAACGCTGGT GCGCCTGAA GTGCAACACT 1620
 TGTATTTGTC ACAAGGATCA CATGACCTG CGAACCCATG TGAACGCCAT CGCCCTGCAA 1680
 25 GGCTGTGAGG TGCCCGCGGG CTTTGGGCCC CGACCCAT TTGCTTCAG GATCCTGCGC 1740
 AACCGCAGG AGGTGGCCAT CTTGAGGCA AGCTGTTCAG AGGACATGG TCGCTGGCTC 1800
 GGCCTGCTGC TGTGTGAGAT GGCCTCCAGA GTCACTCGG AGCGCTGCA CTATGACTAC 1860
 GTGATGTGAG AGACCTTAAC CAGCATCGTC AGTGCTGGG GCAACTCCTT CCTATATGCA 1920
 AGATCCTGCC AGAATCAGTG GCTGAGCCC CGAGTCTATG ATGATGTTC TTATGAAAG 1980
 30 ATGACGAGG AGGAGCCCGA GCGCCCRACA GGGGCCAGG TGAAGCTCA CGCTCCTCC 2040
 TGCAGTGAAG AGTCCCATCG TGTGACCCG CAGGTCAAAG TCAACGCCA CGCTCCAGT 2100
 GGCATCAAT ACAAGTATGG CAAGAACCBA GCGGAGGAG ATGCCCGAG GTACTTGGTA 2160
 GAAAGAGAGA AGCTGGAGAA AGAGAAAGAG ACATTCGGA CAGAGCTGAT AGCACTGAGA 2220
 CAGGAGAGA GGAAGCCATT CGAGCAGCC CAGGAGCAA ATTAAGGCT 2280
 35 CTGGAAGAG CCGTGGCCAC CCGGAAGCT CAGTGTGGG CAAAGGAGGA GCGCCGATT 2340
 GACCTGAGC TGAAGCTGGT GGCTGTGAAG GAGCGCTGC AGCACTCCTT GGCAGGAGG 2400
 CCAGCCCTCG GGTCTCCGT GAGCAGCAAG CCAAGAGTG GCAACTCTC TGAGGAGAT 2460
 40 ACGCTCACT CCAATGGTGC TCTCTCAGAG AGAATTCTC TGACCTCATC TACACCAGG 2520
 CTTCTCAACC CCAACACTAC TGACATTTTG GACCAGTAA

Seq ID NO: 63 Protein sequence:
 Protein Accession #: fgenes prediction

45 1 11 21 31 41 51
 | | | | |
 MDRQQKRGD DARTCCGAGR ERETGRSEAG EEEGERRAVG RGLRNARRGL GDAALMQRCL 60
 RLPGQPASNQ VQLSEVPQRK LRVPSPSPVA EKVLGHRCL ELLEQLLPRL TGLLSILDHE 120
 50 YLSDTTLEKK MAVASILQSL QPLFAKEVSY LYVNTADLHS GPSFVSLFE EFDCLSDLR 180
 DMPEDDGEPS KGASPELAKS PRLRNAADLP PFLPNKPPE DYEEALPLG PGKSPETISS 240
 HNGCSPSHSI VDGYYEDADS SYPATRVNGE LKSSYNDSDA MSSSYBSYDE EEEEGKSPQP 300
 RHQWPSSEAS MHLVRECRIC AFLLRKKRFG QWAKQLTVIR EDQLLCYKSS KDRQPELRLA 360
 LDTCSIIYVP KDSRHKRHSI RFTQGATEVL VLALQSRQA EEWLKVIREV SKFVGAGBGV 420
 55 EVFRSPVLIC KLDLDKRLSQ EKQTSDDSV GVDNCSITLG RRETCDHKG KKSLLBLKG 480
 SMSRAAGRKI TRIIGFSKIK TLADDLQTS TEEVPCOGY LNVLVNQGWK ERWCRLKNT 540
 LYFHKHMDL RTHVNAIALQ GCEVAPGPG RHFPAPRIIR NRQVAILLEA SCSEDMGRWL 600
 GLLLVEMGRS VTFEALHYD VDVETLSIV SAGRNSFLYA RSCQWQWPEF RVYDDVPYK 660
 MQDEEPRFT GAQVKRHASS CSEKSHRVPD QVKVRHASS ANQYKYGNR ABEADARRYLV 720
 60 EKELEKEKE TIRTELLALR QEKRELKEAI RSSPGAKLKA LEEAVATLEA QCRAKEERRI 780
 DLEKLIVAVK ERLQQLAGG PALGLSVSSK PRSGQLSEED TLTSGALSE RTSLTSSTPG 840
 LNPNTTDL DQ

Seq ID NO: 64 Nucleotide sequence:
 Nucleic Acid Accession #: NM_004126.1
 65 Coding sequence: 108-129 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 | | | | |
 GGCACGAGCT CTGCGCGGCC TTCAGTTGTT TCGGGACGCG CGGAGCTTCG CGGCTCTTCC 60
 AGCGGCTCCG CTGCCAGAGC TAGCCCGAGC CCGGTTCTGG GCGGAAATG CCTGCCCTTC 120
 ACATCGAGA TTTGCCAGAG AAGGAAAAAC TGAATATGGA AGTTGAGCAG CTTGCGAAAG 180
 AAGTGAAGTT GCAGAGACAA CAAGTGTCTA AATGTTCTGA AGAAATAAAG AACTATATTG 240
 75 AAGAAGCTTC TGGAGAGGAT CCTCTAGTAA AGGGAATTC AGAAGACAAG AACCCCTTTA 300
 AAGAAAAAGG CAGCTGTGTT ATTTCAATGA TAACCTGGGA GAACTGCAT CCTAAGTGG 360
 AAGACTAGTT TGTTTAGTT TTCCAGATA AAACCAACAT GCTTTTAAAG GAAGGAAGAA 420

TGAATTAA AGGAGACTTT CTTAAGCACC ATATAGATAG GGTATGTAT AAAAGCATAT 480
 GTGCTACTCA TCTTTGCTCA CTATGCAGTC TTTTAAAGA GAGCAGAGAG TATCAGATGT 540
 ACAATTATGG AATAAGAAC ATTACTTGAG CATGACACTT CTTTCAGTAT ATTGCTTGAT 600
 GCTTCAAATA AAGTTTTC TC

5

Seq ID NO: 65 Protein sequence:
 Protein Accession #: NP_004117

10 1 11 21 31 41 51
 | | | | | |
 MPALHIEDLP EKEKLMKRVB QLRKEVKLQR QQVSKCSREI KNVIBERSGR DPLVKGIPEP 60
 KNPPKKGSC VIS

15

Seq ID NO: 66 Nucleotide sequence:
 Nucleic Acid Accession #: NM_003842.1
 Coding sequence: 1-1236 (underlined sequences correspond to start and stop codons)

20 1 11 21 31 41 51
 | | | | | |
 ATGGAACAAC GGGGACAGAA GGGCCCCGCC GCTTCGGGGG CCGGAAAAG GCACGGCCCA 60
 GGACCCAGGG AGGCGGGGG AGCCAGGCCT GGGCCCCGGG TCCCCAAGAC CCTTGTGCTC 120
 GTTGTCCCGG CGGTCTCTCT GTTGTCTTCA GCTGAGTCTG CTCTGATCAC CCAACAAGAC 180
 25 GTTGTCCCGG AGCAGAGAGC GGGCCCCACA CAAGAAGGT CCAGCCCTC AGAGGGATTG 240
 TGTCCACTCG GACACCATAT CTCAGAAGAC GGTAGAGATT GCATCTCCTG CAAATATGGA 300
 CAGGACTATA GCATCACTG GAATGACCTC CTTTCTGCT TGGCTGCAC CAGGTGTGAT 360
 TCAGGTGAAG TGGAGCTAAG TCCCTGCACC ACGACCAGAA ACACAGTGTG TCAGTGGGAA 420
 30 TGAGGCACCT TCCGGGAAGA AGATTCTCCT GAGATGTGCC GGAAGTGGC CACAGGGTGT 480
 CCCAGAGGGA TGGTCAAGGT CGGTGATGTG ACACCTTGA GTGACATCGA ATGTGTCCAC 540
 AAGAATTCAG GCATCATCAT AGGAGTCACA GTTGCAGCGG TAGTCTTGAT TGTGGCTGTG 600
 TTTGTTTGA AGTCTTTACT GTGGAAGAAA GTCTTCTCT ACCTGAAAGG CATCTGCTCA 660
 GGTGTGTGGT GGGACCTTGA GGTGTGGGAC AGAAGCTCAC AACGACCTGG GGCTGAGGAC 720
 35 AATGTCTCTA ATGAGATCGT GAGTATCTTG CAGCCACCCC AGTCCCTGA GCAGGAATG 780
 GAGTGTCCAG AGCCAGCAGA GCCAACAGGT GTCAACATGT TGTCCCGGG GAGTCAGAG 840
 CATCTGCTGG AACCGGCAGA AGCTGAAAGG TCTCAGAGGA GGAGCTGCT GGTTCAGCA 900
 AATGAAGGTG ATCCCACTGA GACTCTGAGA CAGTGCTTCG ATGACTTTGC AGACTTGGTG 960
 CCTTTGACT CCTGGGAGCC GCTCATGAGG AAGTTGGGCC TCATGGACAA TGAGATAAAG 1020
 40 GTGGCTAAAG CTGAGGCAGC GGGCCACAGG GACACCTTGT ACACGATGCT GATAAAGTGG 1080
 GTCAACAAA CCGGGCGAGA TGCCTCTGTC CACACCCTGC TGGATGCTT GAGAGCGCTG 1140
 GGAGAGAGAC TTCCCAAGCA GAAGATTGAG GACCACTTGT TGAGCTCTGG AAAGTTTCATG 1200
 TATCTAGAAG GTATGCAGA CTCTGCCATG TCCTAA

45 Seq ID NO: 67 Protein sequence:
 Protein Accession #: NP_003833.1

50 1 11 21 31 41 51
 | | | | | |
 MEQRGONAPA ASGARRRHGP GPREARGARP GPRVPKTLVL VVAVLLLVV AESALITQOD 60
 LAPQRAAPQ QKRSPSEGL CPPGHKISE GRDCISCKYG QDVSTHNDL LFLCRLCTCD 120
 SGHVELSPCT TTRNTVCQCE BGTFRBEDSP BMCRKCRGTC PRGMVKVGDG TFWSDIECVH 180
 KESGIIIGVT VAAVVLIVAV FVCKSLMKK VLPYLKIGCS GGGDPERVD RSSQRPQED 240
 NVLNBIVSIL QPTQVPEQM EVQSPAEPTG VMLSPGESE HLLSPAEER SQRRRLVPA 300
 55 NEGDPTEITLR QCFDDPADLV PFDSWEPLMR KLGLMDNEIK VAKAEAGHR DTLTYMLIKW 360
 VNKTRDASV HTLLDALETL GRLAKQKIE DHLLSSGKFM YLEGNADSAM S

60 Seq ID NO: 68 Nucleotide sequence:
 Nucleic Acid Accession #: FGENESH predicted ORF
 Coding sequence: 361- 2220 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51
 | | | | | |
 GGCACCATCT GCTCCCTGCC CTGCCAGAG GGCCTTCAG GACCCAACCTG CTCCAGGAA 60
 TGTGCTGCGG ACAACGGCGG CCTCTGTGAC CGATTCACCTG GGCAGTCCCG CTGCGCTCG 120
 GGTACACTG GGGATCGGTG CCGGAGGAG TGCCCGGTGG GCGCTTTGG GCAGGACTGT 180
 GCTGAGACGT GCGACTGCGC CCGGAGCGCC CGTTGCTTCC CGGCCAACGG CGCATGTCTG 240
 TGGAACACG GCTTCACTGG GACCGCTGC ACGGATCGCC TCTGCCCGA CGGCTTCTAC 300
 70 GGTCTCAGCT GCCAGGCCCT CTGCACCTGC GACCGGAGC ACAGCCTCAG CTGCCACCGG 360
 ATGAACGGGG AGTGCTCCTG CTTGCCGGGC TGGCGGGGCC TCCACTGCA CGAGAGCTGC 420
 CCGCAGGACA CGCATGGGCC AGGTGCCAG GAGCACTGTC TCTGCTGCA CGGTGGCGTC 480
 TCCAGGCTA CCGCGGCTT CTGTCACTGC GCGCGGGTT ACACGGGCC TCACGTGCT 540
 ACTCTTTGTC CTCTGCACAC CTACGCTGTC AACTGTTCTG CACGCTGCTC ATGTGAAAAT 600
 75 GCCATCGCCT GCTCACCCAT CGACGGCGAG TGGCTCTGCA AGGAAGGTTG GCAGCGTGGT 660
 AACTGCTCTG TGGCTGCTCC ACCCGGAACC TGGGGCTTCA GTTGCAATGC CAGCTGCCAG 720
 TGTGCCCATG AGGCAGTCTG CAGCCCCAA ACTGAGCCT GTACCTGCAC CCTGGGTGG 780

5
 10
 15
 20
 25

```

CATGGGCCCC ACTGCCAGCT GCCCTGTCCG AAGGGGCACT TTGGAGAAGG TTGTGCCAGT 840
CGCTGTGACT GTGACCACTC TGATGTGCTGT GACCTGTGTC ATGGAGCTGT TCAGTGCCAG 900
GCTGGCTGGA TGGGTGCCCG CTGCCACCTG TCCTGCCCTG AGGGCTTATG GGGAGTCAAC 960
TGTAGCAACA CCTGCACCTG CAAGAATGGG GGCACCTGTC TCCTTGAGAA TGGCAACTGC 1020
GTGTGTGCAC CGGATTCCCG GGGCCCCCTCC TGCCAGAGAT CCTGTGACCC TGGCCGCTAT 1080
GGCAACGCT GTGTGCCCTG CAAGTGCCTT AACCACTCCT TCTGCCACCC CTGGAACGGG 1140
ACCTGTCTACT GCTGTGCTGG CTGGACAGGC CCGGACTGCT CCGAGGCTG CTCTCTGGGG 1200
ACATTGTGTG CTAAGTCTCT CCAGCCATGC CAGTGTGCTC CTGGAGAAAA GTGCCAOCOA 1260
GAGACTGGGG CTTGTGTATG TCCCCAGGG CACAGTGGTG CACTTGCAG GATTGGAATC 1320
10 CAGGAGCCCT TTAATGTGAT GCGGACCACT CCAATAGCTT ATAATGCTG GGTGTGAGTG 1380
ATTGGCAATG CAGTGTCTGG GTCCCTGTG GTAGCCCTGG TGGCACTGTT CATTGGCTAT 1440
CGGCACTGGC AAAAAGGCAA GGAGCACCACT CACTTGGCTG TGGCTTACAG CAGCGGGGCG 1500
CTGACGCGCT CCGATATGT CATGCCAGAT GTCCCTCCGA GCTACAGTCA CTACTACTCC 1560
AACCCAGCT ACCACACCTT GTGCGAGTGC TCCCAAAACC CCCCACCCCT TAACAGGTT 1620
15 CAGGCCCCGC TCTTTGCCAG CCTGCAGAAC CCTGAGGGGC CAGGTGGGGC CCAAGGGCAT 1680
GATAGCACCA CCACTCTGCC TGCTGACTGG AAGCACCGCC GGGAGCCCCC TCCAGGGCCT 1740
CTGACAGGGG GGAGCAGCGG CTTGGACCGA AGCTACAGCT ATAGCTACAG CAATGGCCCC 1800
GGCCCATCTT ACAATAAAGG GCTCATCTCT GAAGAGGAGC TGGGGGCCAG TGTGGCTTCC 1860
CTGACAGCTG AGAACCCATA TGCCACCATC CGGAGCTGTC CAGCTTGGCC AGGGGGCCCC 1920
20 CCGGAGAGCA GTACATGGA GATGAAAGGC CCTCCTCAG GATCTCCCCC CAGGCAGCCT 1980
CCTCATCTCT GGGACAGCCA GAGCGGGCGG CAACCCAGC CACAGAGAGA CAGTGGCACC 2040
TACGAGCAGC CCGGCCCCCT GATCCATGAC CGAGACTCTG TGGGCTCCCA GCCCCCTCTG 2100
CTCCGGGGCC TACCCCGCGG CCACATGAC TCACCAAGA ACAGCCACAT CCTTGGACAT 2160
25 TATGACTTGC CTCAGTACG GCATCCCCCA TCACCTCCAC TTCGAGCCCA GGACCGTTGA
  
```

Seq ID NO: 69 Protein sequence:

Protein Accession #: FGENESH prediction

30
 35
 40
 45

```

1 11 21 31 41 51
| | | | |
GTICSLPCPE GFHGFNCSQE CRCHNGGLCD RFTQQRCPAP GYTQDRCREE CPVGRPGQDC 60
AETCDAPDA RCFPANGACL CEHGFTRDRC TDRLCDPGFY GLSQAPPTCT DREHSLSCHP 120
MMGRCSCLPQ WAGLHCNESC PQDTHGPGCQ EHCLCLHGGV CQATSGLCQC APGYTGVHCA 180
SICPPDTYGV NCSARCSCEA AIACSPIDGE CVCKBGWRG NCSVPCPPGT WGFSCNASCQ 240
CAHBAVCSBQ TGACTCTPGW HGAHCQLPCP KGQPGEGCAS RCDCHSDGC DPVHGRQCQC 300
AGWMGARCHL SCPEGLMGVN CSMTCTCKNG GTCLPENGNC VCAPGFRGPS CQRSCQPGRY 360
GRRCVPCCKA NNSFCHPSNG TCYCLAGWTG PDCSQRCPFG TFGANCSPQC QCGPGBKCHP 420
PGPLFASLQN HSLAPCRIGI QEPFTVMPTT PVAYNSLGAV IGLAVLSLV VALVALFIGY 480
40 RHWGKQKZKH HLAIVYSSGR LDGSEYVMPD VPPSYSHYYS NPSYHTLSQC SPNPPPPNKV 540
PERPFGAQGH DNHTTLPADW KHRREPPPGP LDRGSSRLDR SYSYSYNSGP 600
GPFYNNGLIS EEBLAGSVAS LSEENFYATI RDLPSLPGGP RESSYMEMKG PPSGSPFRQP 660
PQPFWSQRRR QPQQRDSGT YBQPSPLIHD RDSVGSQPPL PGOLPPGHVD SPKNSHIPGH 720
45 YDLPPVRHPP SPFLRRQDR
  
```

Seq ID NO: 70 Nucleotide sequence:

Nucleic Acid Accession #: NM_005458

Coding sequence: 1..2826 (underlined sequences correspond to start and stop codons)

50
 55
 60
 65
 70
 75

```

1 11 21 31 41 51
| | | | |
ATGGCTTCCC CGGGAGGATC CGGGCAGCCA GGGCGGCGGC GCGCGCGGCC ACGCGCGGCC 60
GCGCGCTGTC TACTGCTACT GCTGTGCTGC CTGCTGCTGC CTCTGCGGCC CCGGCGCTGG 120
55 GCGTGGCGGC GGGCGGCCCC CCGGCGGCGG CCGCAGCAGCC CGCGCTCTCT CATCATGGGC 180
CTCATGCGCG TCACCAAGGA GGTGGCCAAAG GGCAGCATCG GGGCGGTGTG GCTCCCCGCC 240
GTGGAATCGG CCATCGAGCA GATCGCAAC GAGTCACTCC TGCGCCCTTA CTCTCTCGAC 300
CTCGGCTCTC ATGACACGGA GTGCGCAAC GCAAAAGGCT TGAAGCCTT CTACGATGCA 360
60 ATAAATACG GGCAGAACCA CTTGATGGTG TTTGGAGGCG TCTGTCCATC CGTCACATCC 420
ATCATTGCAG AGTCCCTCCA AGGCTGGAAT CTGTGTCAGC TTTCTTTTTC TGCAACCAAG 480
CCTGTTCTAG CCGATAAGAA AAAATACCTT TATTTCTTTC GGACCGTCCC ATCAGACAA 540
GCGGTGAATC CAGCCATTCT GAAATTGCTC AAGCACTACC AGTGAAGACG CGTGGGCAAG 600
GTGACGAAG ACCTTCAGAG GTTCTCTGAG GTGCGGAATG ACCTGACTGG AGTCTGTAT 660
GGCGAGACCA TTGAGATTTC AGACACCGAG AGCTTCTCCA ACATTCCTGT TACCAGTGT 720
65 AAAAAGCTGA AGGGGAATGA TGTGCGGATC ATCTTGGGCC AGTTTGACCA GAATATGGCA 780
GCAAAAGTGT TCTGTGTGTC ATACGAGGAG AACATGTATG GTAGTAAATA TCAGTGGATC 840
ATTCCGGGCT GGTACGAGCC TTCTTGTGTG GAGCAGGTGC ACACGGAAGC CAATCATCC 900
CGCTGCCTCC GGAAGAACTC GCTTGTCTGC ATGGAGGGCT ACATTGCGGT GGATTTGAG 960
70 CCCCTGAGCT CCAAGCAGAT CAAGACCATC TCAGGAAAGA CTCACACGCA GTATGAGAGA 1020
GAGTACAACA ACAAGCGGTC AGGCGTGGGG CCCAGCAAGT TCCACGGGTA CGCTACGAT 1080
GGCATCTGGG TCATGCGCAA GACACTGCAG AGGGCCATGG AGCACTGCA TGCCAGCAGC 1140
CGGACACGAG GGATCCAGGA CTTCAACTAC ACGGACCACA CGTGGGCGAG GATCATCCTC 1200
AATGCCATGA ACGAGACCAA CTTCTTCGGG GTACCGGTTC AAGTTGTATT CCGGAATGGG 1260
75 GAGAGAAATG GAGCAATTAA ATTTACTCAA TTTCAAGACA GCAGGAGGAT GAAGGTGGGA 1320
GAGTACACG CTGTGGCCGA CACACTGGAG ATCATCAATG ACACCATCAG GTTCCAGGA 1380
TCCGAACCA CAAAAGACAA GACCATCATC CTGGAGCAGC TCCGGAAGAT CTCCCTACCT 1440
  
```

5
 10
 15
 20
 25

```

CTCTACAGCA TCCTCTCTGC CCTCACCATC CTCGGGATGA TCATGGCCAG TGCTTTTCTC 1500
TTCTTCAACA TCAGAAACCG GAATCAGAAG CTCATAAAGA TGTGAGTCC ATACATGAAC 1560
AACCTTATCA TCCTTGGAGG GATGCTCTCC TATGCTTCCA TATTTCCTT TGCCCTTGAT 1620
GGATCTCTTG TCCTGAAAAA GACCTTTGAA ACACCTTGCA CGCTCAGGAC CTGGATTCTC 1680
ACCGTGGGCT ACACGACCGC TTTTGGGGCC ATGTTTGCAA AGACCTGGAG AGTCCACGCC 1740
ATCTTCAAAA ATGTGAAAAA GAAGAAGAAG ATCATCAAGG ACCAGAAACT GCTTGTGATC 1800
GTGGGGGGCA TGCTGCTGAT CGACCTGTGT ATCTGATCT GCTGGCAGGC TGTGGACCCC 1860
CTGCGAAGGA CAGTGGAGAA GTACAGCATG GAGCGGAGC CAGCAGGAGC GGATATCTCC 1920
ATCGCCCTC TCCTGGAGCA CTGTGAGAAC ACCCATATGA CCATCTGGCT TGGCATCGTC 1980
TATGCCCTACA AGGACTTCT CATGTTGTTC GGTGTTTCT TAGCTTGGGA GACCCGCAAC 2040
GTCAGCATCC CCGCACTCAA CGACAGCAAG TACATCGGGA TGAGTGTCTA CAAAGTGGGG 2100
ATCATGTGCA TCATCGGGG CGCTGTCTCC TTCTGACCC GGGACCAACC CAATGTGCAG 2160
TTCTGCTATG TGGCTCTGGT CATCATCTTC TGCGAGACCA TCACCTCTG CTGTGTATTC 2220
GTGCGAAGC TCATCACCTT GAGAACAAAC CCAGATGCAG CAACGCAGAA CAGGCGATTC 2280
CAGTTCACTC AGAATCAGAA GAAAGAAGAT TCTAAAACT CCACCTCGGT CACCAAGTGTG 2340
AACCAAGCCA GCACATCCCG CCTGAGGGGC CTACAGTCAG AAAACCATCG CCTGGAATG 2400
AAATGTGCA AGCTGGATAA AGACTTGGAA GAGGTCAACA TGCAGCTGCA GGACACACCA 2460
GAAAGACCA CTACATTAA ACAGAACCA TACCAAGAGC TCAATGACAT CCTCAACCTG 2520
GGAAACTTCA CTGAGGCAC AGATGGAGGA AAGGCCATTT TAAAAAATA CTCTGATCAA 2580
AATCCCGAGC TACGTGGAA CACAACAGAG CCCTCTCGAA CATGCAAGA TCCTATAGAA 2640
GATATAACT CTCCAGAAC CATCCAGCT CGGCTGTCCC TCCAGCTCCC CATCTCCAC 2700
CACGCTACG TCCATCCAT CGGAGGCGTG GACGCCAGCT GTGTGAGCCC CTGCTCAGC 2760
CCCACGCCA GCCCCGCCA CAGACATGTG CCACCTCTCT TCGAGTCAT GGTCTGGGC 2820
CTGTAA
  
```

Seq ID NO: 71 Protein sequence:
 Protein Accession #: NP_005449

30
 35
 40
 45

```

1 11 21 31 41 51
MASPRRSGQP GRPPPPPPPP ARLLLLLLLL LLLPLAPGAM GWARGAPRPP PSSPPLSIMG 60
LMPLTKKVAK GSIGRGVLPV VELAISEQIRN ESLRPPYFLD LRLYDTBCDN AKGLKAFYDA 120
IKYGPNNHLMV FGGVCPSEVTS IIAESLQGMN LVQLSPAATT PVLADKKKYP YFFRTVPSDN 180
AVNPAILKLKL KHYQKRVGT LTQDVQRFSE VRNDLTGVLY GEDIEISDTE SPSPNDPCTSV 240
KKLKGNDVRI ILGQFDQNM A KVFCAYEE NMYGSKYQWI IPGWYEPBWM EQVHTBANSS 300
RCLRNKLLAA MSYIGVDFF PLSSKQIKTI SKKTPQQYER EYNNKRSVGW PSKPHGYAYD 360
GIWVIKTLQ RAMETLHASS RHQRIQDFNY TDHTLGRILL NAMNBTNFFG VTGQVVFRRG 420
ERMOTIKFTQ FQDSREVKVG EYNVADTLE IINDTIRFQG SEPPKDKTII LRQLKISLP 480
LYSILSALT LGMMSASAPL PFNIKRNQK LKMSSEPYMN NLIYLGMLSL YASIFLPLGLD 540
GATVATKTFE TLCTVRTWIL TVGYTTAFGA MFAKTWRVHA IFKNVKKKKK IKDKQLLVI 600
VOGMLLIDIC ILICWQAVDP LRRTVRYKYM EPDPAGRDIS IRPLLEHCEN THMTIWLQIV 660
YAYIKLMLF GCPLAMETRN VSIPALNDSK YIGMSVINVG IMCIIGAAVS FLTRDQPNVQ 720
FTATVAVIIF CSTITLCLVF VPKLITLRTN PDAATQNRFP QFTQNKQKED SKTSTSVTSV 780
NQASTSRLLG LQSNHRLRM KITELDKDLE EVTMLQDTP EKTYYIKQNH YQELNDILNL 840
GNFTSTDDG KAILKHLQ NPQLQWNTS PSRTCKDPIE DINSPEHIQR RLSLQLPILH 900
HAYLPSIGGV DASCVSPVUS PTASPRHRHV PPSFRVMVSG L
  
```

Seq ID NO: 72 Nucleotide sequence:
 Nucleic Acid Accession #: NM_005795
 Coding sequence: 522-1940 (underlined sequences correspond to start and stop codons)

55
 60
 65
 70
 75

```

1 11 21 31 41 51
GCACGAGGGA ACAACCTCTC TCCTCTSCAGC AGAGAGTGTG ACCTCTGTCT TTAGGACCAT 60
CAAGCTCTGC TAACCTGAAT TCATCCTAAT TGCAGGATCA CATTGCAAGC CTTTCACTCT 120
TTCCCACTTT GCTTGTGGGT AAATCTCTTC TGCGGAATCT CAGAAAGTAA AGTTCCATCC 180
TGAGAAATTT TCACAAAGAA TTTCCTTAAG AGCTGGACTG GGTCTTGACC CCTGGAATTT 240
AAGAAATTTCT TAAAGACAAT GTCAATATG ATCCAAGAGA AAATGTGATT TGAGTCTGGA 300
GACAATTTGT CATATCTCTT AATAATAAAA ACCCATACTA GCCTATAGAA AACATATTT 360
GAATAATAAAA AACCCATACT AGCCTATAGA AAACAATATT TGAAAGATTG CTACCACTAA 420
AAGAAAACT ACTACAATT GACAGAAGCTG CTGCAAACTT CAATTGGTCA CCACAACCTG 480
ACAAGGTTGC TATAAAACAA ATTTGCTACA ACTTCTAGTT TATGTTATAC AGCATATTTT 540
ATTTGGGCTT AATGATGGAG AAAAAGTGA CCCTGTATTT TCTGTTCTC TTGCTTTT 600
TTATGATTCT TGTACAGCA GAATTAGAAG AGAGTCTCTG GAGCTCAATT CAGTGGGAG 660
TTACTAGAAA TAAATCATG ACAGCTCAAT ATGAATGTTA CCAAAGATT ATGCAAGACC 720
CCATTCAACA AGCAGAAGGC GTTACTGCA ACAGAACCTG GGATGGATGG CTCTGCTGGA 780
ACGATGTTGC AGCAGGAAT GAATCAATGC AGCTCTGCCG TGATTACTTT CAGGACTTTG 840
ATCCATCAGA AAAAGTTACA AAGATCTGTG ACCAAGATGG AAATGTTT AGACATCCAG 900
CAAGCAACAG AACATGGACA AATTATACCC AGTGTAAATG TAACACCCAC GAGAAAGTGA 960
AGACTGCACAT AAATTGTTT TACCTGACCA TAATTGGACA CGGATTGTCT ATTGTCATAC 1020
TGCTTAATCT GCTTGGCATA TTCTTTTATT TCAAGGCCT AAGTTGCCAA AGGATTACCT 1080
TACACAAAAA TCTGTCTTC TCATTGTGTT GTAACCTGTG TGTAAACATC ATTCACTCA 1140
CTGCAAGTGC CAACAACAG GCCTTAGTAG CCACAATCC TGTAGTTGC AAAGTGTCCC 1200
ATGTCATTTCA TCTTACCTG ATGGGCTGTA ATTACTTTTG GATGCTCTGT GAAAGCATTT 1260
ACCTACACAC ACTCATTTGT GTGGCCGTGT TTGCAGAGAA GCAACATTTA ATGTGGTATT 1320
  
```

5
 10
 15
 20
 25
 30
 35
 40
 45

ATTTCTTGG CTGGGGATT CCACCTGATT CTGCTTGTAT ACATGCCATT GCTAGAAGCT 1380
 TATATTACAA TGACAAITGC TGGATCAGTT CTGATACCCA TCTCCTCTAC ATTATCCATG 1440
 GCCCAATTGG TGTGCTTTA CTGGTGAATC TTTTTCCTT GTTAAATATT GTACGGGTTC 1500
 TCATCACCAA GTTAAAGTT ACACACCAAG CGGAATCCAA TCTGTACATG AAAGCTGTGA 1560
 GAGCTACTCT TATCTTGGTG CCATTGCTTG GCATTGAATT TGTGCTGATT CCATGGCGAC 1620
 CTGAAGGAAA GATTGACAGG GAGGTATATG ACTACATCAT GCACATCCTT ATGCACTTCC 1680
 AGGGTCTTTT GGTCTCTACC ATTTCTCTGT TCTTTAATGG AGAGGTTCAA GCAATTCTGA 1740
 GAAGAACTTG GAATCAATAC AAAATCCAAT TTGGAAACAG CTTTCCAAC TCAGAAAGCTC 1800
 TTGCTAGTGC GTCTTACACA GTGTCAACAA TCAGTGTGGG TCCAGGTTAT AGTCATGACT 1860
 GTCTCTAGTG ACACCTAAAT GSAAAAAGCA TCCATGATAT TGAAAATGTT CTCTTAAAC 1920
 CAGAAAATT ATATAATTG AAAATGAAGG ATGGTTGTCT CACTGTTTGG TGCTTCTCCT 1980
 AACTCAAGGA CTGGACCCA TGACTCTGTA GCCAGAAGAC TTCAATATTA AATGACTTTG 2040
 GGGAAATGTA TAAAGAAGAG CCTTCACATG AAATTAGTAG TGTGTGATA AGAGTGTAAAC 2100
 CTTCCAGCTCT ATGTGGGAAA AAGAAATCC TGGTGTGTAA TGTGTGTGAG TAAATCTCC 2160
 CACTATGCTG GATGTGACGC TACTAACCTG ACATCACCAG GTGTGGAAAT GGAGAAAGC 2220
 ACAATCAACT TTTCTGAGCT GGTGTAGGCC AGTTCACGCA CACCAATTGAT GAATCAAAAC 2280
 AAATGGCTGT AAAACTAAAC ATACATGTGG GGCATGATTC TACCCTTATT CCCCCAAGA 2340
 GACCTAGTGA AGGTCTATAA ACATGAAGGG AAAATTAGCT TTTAGTTTTA AAACCTTTTA 2400
 TCCCATCTTG ATGGGGCAG TTGACTTTTT TTTTTCCTCA GAGTGGCGTA GTCTTTTGG 2460
 TAACCTACCT CTCAAATGGA CAATACCAGA AGTGAATTAT CCTGTCTGGC TTTCTTTTCT 2520
 CTATGAAAAG CAACCTAGTA CAATTGTTAT GATCTACTCA TTTGCTGACA CATCAGTTAT 2580
 ATCTTGTGGC ATATCCATTG TGGAACTGG ATGAACAGGA TGTATAATAT GCAATCTTAC 2640
 TTCTATATCA TTAGGAAAAC ATCTTAGTTG ATGCTACAAA ACACCTTGTC AACCTCTTCC 2700
 TGCTCTACCA AACAGTGGGA GGGAAATCCT AGCTGTAAAT ATAAATTTTG AGAATTTCC 2760
 TTTCTATGTA TAAACAAATT AGCAATCAAT TTATATAAAG AAAATCAATG AAGGATTTCT 2820
 TATTTCTCTG GAATTTTCTG AAAAGAAATT GTGAAAATG AGCTGTGAAA TACTCCATTA 2880
 TTTTATTTTA TAGTCTCAAA TCAATATCAT ACAACCTATG TAAATTTTAA AGCAAAATA 2940
 TAAATGCAACA ATGTGTGTAT GTTAATATCT GATACGTGAT CTGGGCTGAT TTTTAAATA 3000
 AAAATAGAGTC TGGAAATGTA TATTTGTGTA ATATTTTAAA GACAAACAGA TGCCAGCATC 3060
 AGAAGTCTGT TTGAGACTA AGAGAACAGA AACATCTATC ATAAGATATA TTTATTTTAA 3120
 AAACACAAAG TCACTATTTT ACTGAATATA TTTGTTTGA TAACTCATAC CTTAATAATA 3180
 GGTGTGTTGG ACATATTTCT TTTTCAATTT TGACAATGAA CTCACATTCT AATCCAGAAA 3240
 TTTTAAACAA CTACTGTGAT AAATACCAAT CTGCTACTTT TATAGATTIT ACCCCATTAA 3300
 AATATTACTT TACTGACTTT TACTATGTGA AGATATATAG CTTTGGAAAT GTCCAGGCT 3360
 ATTTCAAGAA TATAAAAAAC TAGAAGGATA CTATATATAC CATATACAAT GCTTTAATAT 3420
 TTTAATAGAG CTACTGTATA TAATACAAAT TAGGGAAATA CTTGAATATA TCATTGAGAA 3480
 AAAATTTATG TCAGATCTTA CTGAATTTAT GTGAGACTTT ATTAAATAAA GATAGAAGAA 3540
 AACCTTGCTA ATGAATTTAA GTGAAATTTG CATGGGATTC AGTTTCTCTA ATGTATTTT 3600
 CGGCTGAGTA TCTTAAAGAA CAAGAATGAC TTCAATTAGT AAAAGTCAAT TTTGGGAAA 3660
 GTCATGGGTA TCTGTTTTT AAGTGTGTCA ATCTGATTA AATGGATGAA ACAAAATACT 3720
 CATGATAAGT TGTTCCTTAA GCTGTCAATA TGTCAATAGA TGGTCAAGTT AGAAGTATT 3780
 TCAAAATGCT AAGACAAATT ATCTAAATTC GTAAGAATTA ACATATAGAA TGGTCTGGTC 3840
 AGTACATTTA TAATTTATCT ATGCATGAAA AAGTATTGTT TGTGTTGAAA CATGAATTTT 3900
 ATAGCAAGCT GCCATAGAAA GGA

Seq ID NO: 73 Protein sequence
 Protein Accession #: NM_005795

50
 55
 60

1 11 21 31 41 51
 MLYSIFHLGL MMEKKCTLYF LVLLPFFMIL VTAELESPE DBIQLGVTRN KIMTAQYECY 60
 QKIQMDPIQQ AEGVYCNRTW DQWLQNDVA AGTBSMLCP DYPQDFDPE KVKICDQDG 120
 NWFRHPASNR TWNYTQCNV BTHEKVTAL NLFYLTIGH GLSIASLLIS LGIFYPFKSL 180
 SCQRITLHKN LFFSPVCNSV VTIHLTAVA NNQALVATNP VSKVSPFIH LYLNGCNYFW 240
 MLCBGIYLT LIVAFAPEK QHLMYYFLG WGFPLIPACI HALARSLYIN DNCWISSDTH 300
 LLYIHPGIC AALLVNLFFL LNIWRVLTK LKVTQAESN LYMKAVRATL ILVPLLGIEF 360
 VLIWPRFEGK IABEVDYIN HILMHFQGLL VSTIFCFPNG EVQALLRRNW NQYKIQFGNS 420
 FSNSEALRSA SYTVSTISDG POYSHDCPSB HLNGKSIHDI ENVLKPENL YN

Seq ID NO: 74 Nucleotide sequence
 Nucleic Acid Accession #: NM_000450.1
 Coding sequence: 117..1949 (underlined sequences correspond to start and stop codons)

65
 70
 75

1 11 21 31 41 51
 CCTGAGACAG AGGCAGCAGT GATACCCACC TGAGAGATCC TGTGTTGAA CAACGTCTTC 60
 CCAAAACGGA AAGTATTTC AAGCCTAAAC TTTGGGTGAA AAGAACTCTT GAAGTCATGA 120
 TTGCTTACAA GTTCTCTCA GCTCTCACTT TGGTGTCTCT CATTAAAGAG AGTGGAGCCT 180
 GGTCTTACAA CACCTCCAAG GAAGCTATGA CTTATGATGA GGCCAGTGCT TATTGTGAGC 240
 AAGGTACAC ACACCTGGTT GCAATTCAAA ACAAAGAAGA GATGTAGTAC CTAACCTCCA 300
 TTTAGAGCTA TTCAACAAGT TATTACTGGA TTGGAATCAG AAAAGTCAAC AATGTGTGGG 360
 TCTGGGTAGG AACCCAGAAA CCTCTGACAG AAGAAAGCCA GAACGTGGCT CCAAGTGAAC 420
 CCAACATAG GCAAAAAGAT GAGGACTCGG TGGAGATCTA CATCAAGAGA GAAAAGATG 480
 TGGCATGTG GAATGATGAG AGGTGCAGCA AGAAGAAGCT TGCCCTATGC TACACAGCTG 540
 CCGTACCAA TACATCTGCG AGTGGCCACG GTGAATGTGT AGAGACCATC AATAATTACA 600

5 CTTCGAAGTG TGACCCCTGGC TTCAGTGGAC TCAAGTGTGA GCAAATTGTG AACTGTACAG 660
 CCTGTGAATC CCTGTAGCAT GGAAGCCTGG TTTGCAGTCA CCCACTGGGA AACTTCAGCT 720
 ACNATTCCTC CTGCTCTATC AGCTGTGATA GGGGTTACCT GCCAAGCAGC ATGAGAGCCA 780
 TGCAGTGAT GTCTCTGGG GAATGGAGTG CTCCTATTCC AGCCTGCAAT GTGGTTGAGT 840
 GTGATGCTGT GACCAATCCA GCCAATGGGT TCGTGAATG TTTCCAAAAC CCTGGAAGCT 900
 TCCCATGGAA CACAACCTGT ACATTGACT GTGAAGAAGG AITTGAACTA ATGGGAGCCC 960
 AGAGCCTTCA GTGTACCTCA TCTGGGAATT GGGACACGA GAAGCCAACG TGTAAAGCTG 1020
 TGACATGCAG GGCCTGCOGC CAGCCTCAGA ATGGCTCTGT GAGGTGCAGC CATTCOCCTG 1080
 10 CTGGAGAGTT CACCTTCAAA TCATCCTGCA ACTTCACCTG TGAGGAAGGC TTCACTGTTC 1140
 AGGGAACGAG CCAGGTTGAA TGCACCACTC AAGGOCAGTG GACACAGCAA ATCCAGTTT 1200
 GTGAAGCTTT CCAAGTGCACA GCCTTGTCCA ACCCCGAGCG AGGCTACATG AATTGTCTTC 1260
 CTAGTGTCTC TGGCAGTTTC CGTTATGGGT CCAGCTGTGA GTTCTCTGTG GAGCAGGGTT 1320
 TTGTGTGAA GGGATCCAAA AGGCTCCAAT GTGGCCCCAC AGGGGAGTGG GACAAOGAGA 1380
 AGCCCACTAG TGAAGCTGTG AGATGCGATG CTGTCCACCA GCCCCGAGG GGTTTGGTGA 1440
 15 GGTGTGCTCA TTCCCTATT GGAGAATTCA CCTACAAGTC CTCTGTGCCC TTCAGCTGTG 1500
 AGAGGGGATT TGAATTATAT GGAATCAACT AACTTGAGTG CACATCTCAG GGACAAATGGA 1560
 CAGAAAGTAT TCCTCTCTGC CAAGTGGTAA AATGTTCAAG CCTGGCAGTT COGGGAAGA 1620
 TCACATGAG CTGCAAGTGG GAGCCCGTGT TTGGCACTGT GTGCAAGTTC GCCTGTCTCTG 1680
 20 AAGATGGAG GCTCAATGGC TCTGCAGCTC GGACATGTGG AGCCACAGGA CACTGTGTCTG 1740
 GCCTGTCTACC TACCTGTGAA GCTCCCACTG AGTCCAACAT TCCTTGTGTA GCTGTGACTTT 1800
 CTGCTGCTGG ACTCTCCCTC CTGACATTAG CACCATTCTT CCTCTGGCTT CGGAAATGCT 1860
 TACGGAAGAC AAGAAATTTT GTTCTGTCCA CGAGCTGCCA AAGCCTTGAA TCAGACGGAA 1920
 GGTACCAAAA GCCTTCTTAC ATCCTTTAAG TTCAAAAGAA TCAGAAACAG GTGCATCTGG 1980
 25 GGAAGTAGAG GGATACACTG AAGTTAACAG AGACAGATAA CTCTCTCGG GTCTCTGGCC 2040
 CTCTCTGCCT ACTATGCCAG ATGCCTTTAT GGCTGAAACC GCAACACCCA TCACCACTTC 2100
 AATAGATCAA AGTCCAGCAG GCAAGGACGG CCTTCAACTG AAAAGACTCA GTGTTCCTTT 2160
 TCCTACTCTC AGGATCAAGA AAGTGTGGC TAAAGAGGG AAGGATATT TTCTTCCAAG 2220
 CAAAGGTGAA GAGACCAAGA CTCTGAAATC TCAGAATTCC TTTCTTAAC CTCCCTGTCT 2280
 30 CGCTGTAAAA TCTTGGCACA GAAACACAAT ATTTGTGGC TTTCTTTCTT TTGCCCTTCA 2340
 CAGTGTTCGG ACAGCTGATT ACACAGTTGC TGTCTAAGA ATGAATATA ATTATCCAGA 2400
 GTTTAGAGGA AAAAATGAC TAAAAATATT ATAACTTAAA AAAATGACAG ATGTTGAATG 2460
 CCCACAGGCA AATGCATGGA GGGTGTGTAA TGGTGCAAT CTACTGAAT GCTCTGTGG 2520
 AGGGTACTA TGCACAATTT AATCACTTTC ATCCCTATGG GATTCAGTGC TTCTTAAAGA 2580
 35 GTTCTTAAGG ATTTGTATAT TTTTACTTGC ATTTGAATATA TTATAATCTT CCATCTTCT 2640
 TCAATCAATA CAAGTGTGGT AGGGACTTAA AAAACTTGTA AATGCTGTCA ACTATGATAT 2700
 GGTAAAGATT ACTTATTCTA GATTACCCCC TCATTGTTTA TTAACAAAT ATGTTACATC 2760
 TGTTTTAAAT TTATTTCAAA AAGGGAAACT ATTTGCCCT AGCAAGGCAT GATGTTAAAC 2820
 AGAATAAAGT TCTGAGTGT TTTACTACAG TTGTTTTTGG AAAACATGGT AGAATTTGAG 2880
 40 TTTAAGAAAT GAATGGAAGG TTTGTATATT GTGAGATATT TTTTCAGAAA TATGTGGT 2940
 CCACGATGAA AAACTTCCAT GAGGCCAAAC GTTTTGAAC TAAATAAGCA TAAATGCAAA 3000
 TCAACAAAGG TATAATTTTA TGAATGTCTT TGTGGAATAA GAATACAGAA AGATGGATGT 3060
 GCTTGTGATT CCTACAAAGA TGTTTGTGAG ATGTGATATG TAAACATAAT TCTGTATAT 3120
 TATGGAAGAT TTTAAATICA CAATAGAAAC TCACCATGTA AAAGAGTCAT CTGTGATATT 3180
 45 TTTAAGCAAT GAAGATGTCT AATAGTTATT CCTATTGTGT TTTCTCTGT ATGTTAGGGT 3240
 GCTCTGGAAG AGAGGAATGC CTGTGTGAGC AAGCATTAT GTTTATTAT AAGCAGATT 3300
 AACATTTCCA AAGGAATCTC CAGTTTTCAG TTGATCACTG GCAATGAAAA ATTCTCAGTC 3360
 AGTAATTTGCC AAGAGCTGCTC TAGCCTTGAG GAGTGTGAGA ATCAAACTC TCCTACACTT 3420
 50 CACTTAACCT AGCATGTGTT GAAAAAATAA GTTTCAGAGA AGTTCTGGCT GAACACTGGC 3480
 AACGACAAAG CCAACAGTCA AAACAGAGAT GTGTAAGGA TCAGAACAGC AGAGGTTCTT 3540
 TTAAGGGGCG AGAAAACTC TGGGAAATAA GAGAGAACAA CTACTGTGAT CAGGCTATGT 3600
 ATGGAATACA GTGTTATTTT CTTTGAATTT GTTTAAGTGT TGTAAATATT TATGTAACT 3660
 GCATTAGAAA TTAGCTGTGT GAAATACCAG TGTGGTTTGT GTTTGAGTTT TATGTAGAA 3720
 55 TTTAAATTAT AACTTAAAT ATTTTATAAT TTTTAAAGTA TATATTTATT TAAGCTTATG 3780
 TCAGACCTAT TTGACATAAC ACTATAAAGG TTGACAATAA ATGTGCTTAT GTTT

Seq ID NO: 75 Protein sequence:

Protein Accession #: NP_000441

60 1 11 21 31 41 51
 MLASQFLSAL TLVLLIKESG AWSYNTSTEA MYDEASAYC QQRVTHLVAI QNKEEIEYLN 60
 SILGYSPSYA WIGIRKVMNV WVVVGTQKPL TEEAKWAPG EPNNRQKDED CVRIYIKREK 120
 DVMWNDERC SKKKLALCYT AACTNTPSCG HGECVETIMN YTCCKDPGFS GLKCSQIVNC 180
 65 TALESPERGK LVCSHPLGNF SYNSSCSISC DRGYLPSSME TMQCHSSGSEW SAIPACNVV 240
 ECDAVTNPAN GFVECPQNPQ SPPWNITCTP DCBEGFELMG AQLQCTSSG NWDNEKPTCK 300
 AVTCTRAVRP QNGSVRCSHS PAGEFTFKSS CNFTCRBGM LQGPQVVECT TQGMWTTQIP 360
 VCEAFQCTAL SNPERGYMNC LPSASGSPRY GSSCEPSCQE GFVLKGSRLK QCGPTGEWDN 420
 EKPTCEAVRC DAVHQPPLGL VRCASPIGE PTYKSSCAFS CEEGFELYGS TQLECTSGQG 480
 70 WTEVPSCQV VKCSLAVPG KINMSCSGEP VFGTVCKFAC PSQWTLNGSA ARTCGATGHW 540
 SGLPPTCEAP TESNIPLVAG LSAAGLSLLT LAPFLWLRLK CLRKAKKFVP ASSCQSLESD 600
 GSYQKPSYIL

Seq ID NO: 76 Nucleotide sequence:

Nucleic Acid Accession #: NM_031439

Coding sequence: 69..1235 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
5 | | | | | |
CCCGACCCGT GCGAGGGCCA GGTCCGCGCC TGCCCCGCCA GCGAAGCGA GCGGACCCGC 60
GTGGCGCCGT GSCCTCGCTG CTGGGAGCCT ACCCTTGGCC CGAGGGTCTC GAGTGCCCGG 120
CCCTGGACCG CGAGCTGTGG GATGACAAAT CGCGCGCGCC CGTCCCGCGG CCCCOCGGGG 180
ACCAAGGGCTC CGAGAGCCGT ATCCGCGCGC CCATGAACGC CTTCATGGTT TGGGCCAAGG 240
ACGAGAGGAA GTGCTGGAAG GTGCAGAACG CGGACCTGCA CAACGCGGAG CTCAGCAAGA 300
10 TGCTGGGAAA GTGCTGGAAG GCGCTGACGC TGTCCAGAA GAGGCGGTAC GTGGACGAGG 360
CGGAGCGGCT GCGCTGACG CACATGACAG ACTACCCCAA CTACAAGTAC CGGCGCGCGA 420
GGAAGAAGCA GSCCAAGCG CTGTGCAAGC GCGTGGACCC GGGCTTCTCT CTGAGCTCCC 480
TCTCCCGGGA CAGAAACGCC CTGCCGAGAG AGAGAAGCGG CAGCGCGGGG GCGCTGGGGG 540
AGAAGAGGAG CAGGGGTGAG TACTCCCGCG GCACTGCCCT GCCCAGCCTC CGGGGCTGCT 600
15 ACCACGAGGG GCGCGCTGGT GGTGGCGCGG GCGGCACCCC GAGCAGTGTG GACACGTACC 660
CGTACGGGCT GCCACACCT CCTGAAATGT CTCCCTGGA CGTCTGGAG CCGGAGCAGA 720
CCTCTTCTCT CTCCCTCTGC CAGGAGGAGC ATGGCCATCC CCGCGCATC CCCCACCTGC 780
CAGGGCACCC GTACTCACCG GAGTACGCCC CAAGCCCTCT CCACTGTAGC CACCCCTGG 840
20 GCTCCGTGGC CCTTGGCCAG TCCCGCGCGG TCTCCATGAT GTCCCTGTGA CCGCGCTGTC 900
CCCATCTCCG TGCTATTAC TCCCGCGCGA CCTACCAACC ACTCCACTCC AACCTCCAAG 960
CCCACTCTGG CAGCTTTTCC CGGCTCTCTG AGCACCTGG CTTCGACGCC CTGATCAAC 1020
TGAGCCAGGT GGAATCTCTG GGGGACATGG ATCGCAATGA ATTGACCGAG TATTGGAACA 1080
CTCTCGGACA CCCAGACTCC GCCACAGGGG CCATGGCCCT CAGTGGGCA GTTCCGGTCT 1140
25 CCCAGGTGAC ACCAACGGGT CCCACAGAGA CCAGCCTCAT CTCCGTCTTG GCTGATGCCA 1200
CGGCGACGTA CTACAACAGC TACAGTGTGT CATAGAGCTG GAGGCGCCCC GTCCGGCTCAG 1260
CCCTCGCGCC CTCTCTTCT TGTGCTTGA GTGGCAGAGG AGCGTCCAG CCACACCAGC 1320
TTCTCTCCCA CGCTCAGGG CAGGAGGTC TGAAGTGGG CCCCAGAGCC TTGCGCTTAA 1380
GCTGGACTCT CCTTATCOGA GTGCCGCTC TATCCCTTC CCCAGTTC AGCCCCTGCA 1440
30 GCCACATT TAAGTATATT CCTTCAAGTG AGTTTCTCT CAGCCCTGA GAGTTGCTGT 1500
CTCCAGTGG AATGTTCACT GACGCTCTTT CTGTGTAGCC ATCATCGAAA CTAATGGGGG 1560
GACGAGCTG ATAGCCAAGG TCCCTTCTGG TCCAGTTTC TGATTTAGGG TTCTCTCAAG 1620
ATTATAAAG GAAGATGGG AATTTGACT CATTATGAG CTCGCTAACC TACGATCTGG 1680
TGATAATTTT GTGTGCACAG CCCAAGGACC ACGAGGCTTT CTGCACTTTC TGCACCCCTT 1740
TCCAAAGTGA CCACAAATTT TCAAGGGGAC TCATACAATT TGAGAAAAA CAGTCAACCT 1800
35 GATTTGAGAA ATTAACCACT ATGGCTAAT ATATCACAGA AATGGGATT GAGTTAAAC 1860
TATTTTATTT TAAATATACA TTTTAAAGCA GTTCTTTTT TTTGTTAATT TGTTTATTAT 1920
ACACGACATT CAAGAGCCAC CGCGCCGAGC CTACATTAT AATTTTCATT CTCTTTTACC 1980
TATAAAATTC AGTGTATTAG TTTCAATTACA TAGGAGAAAT TATATTTCTA AACATTTTAT 2040
40 GATGTTTAAA AAAAAACAG GCTGTTGTAA AAAAAA AAAA

Seq ID NO: 77 Protein sequence:
Protein Accession #: NP_113627

45 | | | | | |
MASLLGAYFW PEGLECPALD AELSDGQSPF AVPRPPGDKG SESRIRRPMM AFMVWAKDER 60
50 KRLVQNPDL HNAELSKMLG KSKALTLQ KRPYVDEAER LRLQHMODYP NYKYRPRRKK 120
QAKRLCKRVD PGFLSSLSR DQNALPEKRS GSRGALGEKE DRGYSPTGA LPSLRGCYHE 180
GPAGGGGGT PSSVDTPYV LPTPEMSPL DVLEPQTFP SSPQSEHGH PRRIPLPLGH 240
PYSPEYAPST LHCSHPLGSL ALGSPGVSM MSPVPGCPPS PAYYSPATYH PLHSLQLAHL 300
55 GQLSPPEHP GFDALDQLSQ VELLGDMDRN EPDQYLNTPG HPDSATGAMA LSGHVFSQV 360
TFTGPTSETL ISVLADATAT YNYSYSVS

Seq ID NO: 78 Nucleotide sequence:
Nucleic Acid Accession #: XM_035787
Coding sequence: 329..949 (underlined sequences correspond to start and stop codons)

60 | | | | | |
TGCCCCGCCG CGCTCCGAG CGCCCCGGAA GTGATCTGTG GCGGCTGCTG CAGAGCCGCC 60
65 AGGAGGAGGG TGGATCTCCC CAGAGCAAAG CGTCGAGTTC CTCTCTCTCC TTCTCTCTCT 120
CCTCTCTCTC CTCTCTCCAG CGCCGAGGCT CCCCOCGCAC CCGTCAGACT CCTCTCTCGA 180
CGCTCCCGGG CGCGGGGCT TCCAGGCGAC AAGGACGAG TACCTCCGG CCGAGGCCAC 240
GCAGCGCGGG CTTCGAGGC CCTCGGGCGG GCGGACTGGC TCGCGGTGCA GATTCTCTTT 300
AATCCTTTGG TGAAAACGTA GACACAAAT GCGTGCAAT AAGCCCAAGG GTCAGATTTC 360
70 TTTGGCTTTA CACAAAGTCA TCATGGTGGG CAGTGGTGGC GTGGGCAAGT CAGCTCTGAC 420
TCTACAGTTC ATGTACGATG AGTTTGTGGA GGACTATGAG CCTACCAAAG CAGACAGCTA 480
TCGGAAGAAG GTAGTGCTAG ATGGGGAGGA AGTCCAGATC GATATCTTAG ATACAGCTGG 540
GCAGGAGGAC TACGCTGCAA TTAGAGACAA CTACTTCOGA AGTGGGAGG GGTCTCTCTG 600
75 TGTCTTCTCT ATTACAGAAA TGGAATCCTT TGCAGCTACA GCTGACTTCA GGGAGCAGAT 660
TTTAAAGATA AAGAAGATG AGAATGTTCC ATTTCTACTG GTTGTAACA AATCAGATT 720
AGAAGATAAA AGACAGGTTT CTGTAGAGAG GGCAAAAAC AGAGCTGAGC AGTGAATGT 780
TAACATCGTG GAACATCTG CTAAACACG AGCTAATGTT GACAAGGTAT TTTTGAATTT 840

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75

AATGAGAGAA AATCGAGCGA GAAAGATGGA AGACAGCAAA GAAAAGAATG GAAAAAGAA 900
GAGGAAAGAT TTAGCCAGAG GAATCAGAGA AAGATGCTGC ATTTTATAT CAAAGCCCAA 960
ACTCCTTTCT TATCTTGACC ATACTAATAA ATATAATTIA TAAGCATTGC CATTGAAGGC 1020
TTAATTGACT GAAATTACTT TAACATTTTG GAAATTGTTG TATATCACTA AAAGCATGAA 1080
TTGGAACTGC AATGAAAGTC AAATTACTT TAAAAAGAAA TTAATATGGC TTCACCAAGA 1140
AGCAAAAGTC AACCTATTTC ATAAITGCTT ACATTATCA TGGTCCTGAA TGTAGCGTGT 1200
AAGCTTGTGT TTCTTGGGCA GTCTTTCTTG AAATGAAGA GGTGAAATGG GGTGGGGAG 1260
TGGGAGGAAA GGTGACTTCC TCTGGTGTG ATTATAAGC TTAAATTTTA TATCATTTTA 1320
AAATGTCTTG GTCTTCTACT GCCTTGAAAA ATGACAAATG TGAACATGAT AGTTAACTA 1380
CCACTTTTTT TAACCATTAI TATGCAAAAT TTAGAAGAAA AGTTATTGGC ATGGTTGTG 1440
CATATAGTTA AACTGAGAGT AATTCATCTG TGAATCTGCT TTAATTACCT GGTAGTAAAC 1500
TTAGAAAAGT GGTGTAAACT TGTACATGGA ATTTTGTGAA TATGCTCTAA TTTAGAAACT 1560
GAAAAATATC TGGTTATATC ATTCTGGGTG TGTCTTACT GACACCAGGG GTCCGTGACC 1620
CCATGTGCTC TGGTGAGAAA ATATATGCTT GGCACAGCTT TTGTATAGAA AATTCTTGAG 1680
AAGTAATCTG CCGCTAGAAG TCTGTCCAAA TTTAAATGT GTGCCATATT CTGGTTCTTG 1740
AAATAGATTC TCCAGAGCTC TTTGATGCT TTTAATAAAC TGCAAGTTCA TTTTAAATGA 1800
AGGCCAGCA TATATACTTG CAAGATAATT TTCAGCTGCA AGGATTCAGC ACCAGTTATG 1860
TTCTGAATGA CCTCTCTTTT CTCTGAGATT CTGTCCTCTG GAAATCCCTT TCTGTAGTGT 1920
GTGAGCATGT AAGTGTAAAG TTTTAACTCT GGGAGCAGGG CATAGGAAGA AAATGTCACT 1980
AGTCTAATG CATTTTGAC TAGAAGCTT CGGAAAAATA TTATGCTTG CCATCTGTCT 2040
ATTCTTAAAT TTATAATCAT AAAGTTACAG TTGTATACAG GAATTATTAG GAGTAATCT 2100
TTTCTGTTTC TGGTTATAAT GAAGAACACT GTAGCTACAT TTTCAAGAT TAACATCAAG 2160
CCATCAAAAC TGGGTATAGT GCAGAAAACG TGGCACACAC TGACACACAC TTAGGCTGTG 2220
TCACCATTTG TGGGTGATCC TGTGGAAGA ATTCTAGCAT GCTACTTGGG GACATAATTT 2280
CAGTGGGAAA TATGCCACTG ACCGATTTTT TTTTTTCTT CTTTGCACTG GGGCTAGGAC 2340
AGTTGATTC ACAAAAGTAT TTTTTCTTTT TTCTCAGTCC TAATTTGAAC AGGTCAAGA 2400
TGTGTTCAGG CATTCAGGAT AACAGGTGTG TATGTAAAGT TAAAAATAGG CTTTTAGGA 2460
ACTCACTCTT TAGATATTTA CATCCAGCTT CTCATGTAA ATATTGTGCT TTAAGGGT 2520
TGAGATGTAC ATCTTTCATT TCGTATTCTT CATAGGCTAT GCCATGTGCG GAATCAAGT 2580
30 TGAAGATGTA ACACTGGCCA GCGGGCCGAG CAATCTCCAT GTGACTTAT TACAGTCTTA 2640
TTTAAACAGG GTTCTTAACC ACTAACCATT TGACTTTGCT TTGAGACCTT TCCTCTCTCT 2700
GGTACTGAGG TGCTATGAAG CCAACTGACA AAGATGCATC ACCTGTCTTA GGTCTGATGC 2760
ACTACCCGAT TTGTTTATTT GCAATTGAG CCAATTAAAG ACCAATAAAC TTCTTTTTTT

Seq ID NO: 79 Protein sequence:
Protein Accession #: XP_035787

1 11 21 31 41 51
40 MAANKPKGVN SLALHKVIMV GSGGVGKSAL TLQPMYDEFV EDYEPTKADS YRKVKVLDGE 60
EVQIDLLDTA GQEDYAAIRD NYFRSGEGFL CVFSITEMES PAATADFRQG ILRVKEDENV 120
PFLVGNKSD LERKQVSVB EAKNRAEQWN VNYVETSAKT RANVDKVPFD LMRIRARKM 180
BDSREKNGKK KRKSLAKRIR ERCCIL

Seq ID NO: 80 Nucleotide sequence:
Nucleic Acid Accession #: NM_003467
Coding sequence: 89..1147 (underlined sequences correspond to start and stop codons)

50
55
60
65
70
75

1 11 21 31 41 51
GTTTGTGGC TGCGGCAGCA GGTAGCAAAG TGACGCGGAG GGCCTGAGTG CTCAGTAGC 60
CACCGCATCT GGAGAACCAG CGGTTACCAT GAGGGGGATC AGTATATACA CTTCAGATAA 120
CTACACCGAG GAAATGGGCT CAGGGGACTA TGACTCCATG AAGGAACCCCT GTTTCCTGTA 180
AGAAAAATGCT AATTTCATAA AATCTCTCTT GCCCACCATC TACTCCATCA TCTTCTTAAC 240
TGGCATTTGT GGCATAGGAT TGGTCATCTT GGTCTATGGT TACCAGAAGA AACTGAGAAG 300
CATGACGGAC AAGTACAGGC TGCACTCTTC AGTGGCCGAC CTCTCTTTTG TCATCACGCT 360
TCCCTCTCGG GCACTTGATG CCGTGGCAAA CTGCTACTTT GGGAACTTCC TATGCAAGGC 420
AGTCCATGTC ATCTACACAG TCAACCTCTA CAGCAGTGTG CTCATCTGCG CCTTCATCAG 480
TCTGGACCGC TACCTGGCCA TCGTCCACGC CACCAACAGT CAGAGGCCAA GGAAGCTGTT 540
GGCTGAAAAG GTGCTCTATG TTGGCGTCTG GATCCCTGCC CTCTGTCTGA CTATTCCCGA 600
CTTCACTCTT GCCAACGTCA GTGAGGCAGA TGACAGATAT ATCTGTGACC GCTTCTACCC 660
CAATGACTTG TGGGTGGTTG TGTTCAGTTT TCAGCACATC ATGGTTGGCC TTATCTTGCC 720
TGGTATTGTC ATCTGTCTCT GCTATTGCAT TATCATCTCC AAGCTGTAC ACTCCAAGGG 780
CCACCGAAG CGCAAGGCC TCAAGACCAC AGTCATCTCT ATCTGTGCTT TCTTGGCCTG 840
TTGGCTGCTT TACTACATTG GGATCAGCAT CGACTCTTTC ATCTCTCTGG AAATCATCAA 900
GCAAGGGTGT GAGTTTGAGA ACACTGTGCA CAAGTGGATT TCCATCACCG AGGCCCTAGC 960
TTTCTTCAC TGTGTCTGA ACCCATCTT CTATGCTTTC CTGGAGCCA AATTTAAAC 1020
CTCTGCCAG CAGCAGCTCA CCTCTGTGAG CAGAGGGTCC AGCCCTCAAGA TCCTCTCCAA 1080
AGGAAAGCGA GGTGACATT CATCTGTTTC CACTGAGTCT GAGTCTTCAA GTTTTCACTC 1140
CAGCTAACAC AGATGTAAAA GACTTTTTTT TATACGATAA ATAACTTTTT TTTAAGTTAC 1200
ACATTTTTCA GATATAAAG ACTGACCAAT ATTGTACAGT TTTTATTGCT TGTGTGATT 1260
TTGTCTGTGT TTTCTTAGT TTTTGTGAAG TTTAATTGAC TTTTATTAT AAATTTTTTT 1320
TGTTCATAT TGATGTGTGT TTTGTGAGG CTTGTGGCCA AGTTCTTAGT TCGTGTATGT 1380
CTCGTGGTAG GACTGTAGAA AAGGGAACGT AACATTTCCAG AGCGTGTAGT GAATCACGTA 1440
AAGCTAGAAA TGATCCCGAG CTGTTTATGC ATAGATAATC TCTCCATTC CGTGAACGT 1500

TTTCTCTGTT CTTAAGAGCT GATTTTGCTG TAGAAGATGG CACTTATAAC CAAAGCCCAA 1560
AGTGTATATG AATGCTGGT TTTTCAGTTT TCAGGAGTGG GTTGATTCCA GCACCTACAG 1620
TGTACAGCTT TOTATTAAGT TGTAAATAAA AGTACATGTT AAACCTTACTT AGTGTATATG

5 Seq ID NO: 81 Protein sequence:
Protein Accession #: NP_003458

10 1 11 21 31 41 51
MEGISIYTSQ NYTREMGSGD YDSMKPECFR EENANFNKIF LPTIYSIIFL TGIVGNGLVI 60
LVMGVQKKLR SMDKYRLHL SVADLLFVIT LPFMAVDAVA NWYFQNPICK AVHVIYTVNL 120
YSSVLILAFI SLDRYLAIHV ATNSQRPRKL LAKKVYVGV WIPALLLITP DFIFAMVSEA 180
DDRYICDRPY FNDLWVVVPQ PQHIMVGLIL PGIVILSCYC IISKLKSHSK GHQKRKALKT 240
15 TVILILAFPA CWLPHYIGIS IDSPILLRII KQCEPENTV HKWISITEAL APFHCCCLNFI 300
LYAFILAKFK TSAQHALLSV SRGSSILKILS KGRGGHSSV STBSSSSPFI SS

20 Seq ID NO: 82 Nucleotide sequence:
Nucleic Acid Accession #: NM_014959
Coding sequence: 314..1609 (underlined sequences correspond to start and stop codons)

25 1 11 21 31 41 51
CTGGTTCTCA ACTTCTTTTG AAATAATGTT CATAGAGAAG GAGGGCTGTC TGAGATTCTGA 60
GGGAAACAGG CTCTCAGGAC TTCGGGTGCG CATGATGGCT GTGGGCGGTA AACGGGGTTA 120
GTGCAAGCAT CTGGGCCATC TTCCAATGGTA AAAAAGATAC AGTAAAGACA TAAATACCAC 180
ATTTGACAAA TGGAAAAAAA GGAGTGTCCA GAAAAGATA GCAGCAGTGA GGAAGAGCTG 240
CGAGACGGG TATACAGGGA GCTACCCGTG GTTCTGAGA CCTTCTGTA CATCTCACAT 300
30 TTTTCCCAAG AAGATGATGA GACAGAGGCA GAGCCATTAT TGTTCGTCG TGTTCTGAG 360
TGTCACTAT CTGGGGGGGA CATTCCCAGG AGACATTTCG TCAGAGAGGA ATCAATAGT 420
TCTCTTAT GCTTCTAAAG TCTGTTTGA GATGAAAGAA GATTATAAAA ATCTCAGTT 480
TCTGGGCGCT GAAGGAAATG TGGATGTGGA GTTGATTGAT AAGAGCACA ACAGATACAG 540
CCTTGTGTC CCCACTGCTG GCTGTATCT GTGGTCAGCC ACAGGCTCTG GCTTCTGCT 600
35 AAGGGATGAG GTCACAGTGA CGATTGCGTT TGGTTCCTGG AGTCAGCACC TGGCCCTGGA 660
CTCTCAGCAC CATGAACAGT GGCTGTGTGG CGGCCCTTGG TTTGATGTCA CTGCAGAGCC 720
AGAGGAGGCT GTGCCCGAAA TCCACCTCCC CCACTTCATC TCCCTCCAAG GTGAGGTGGA 780
CCTCTCTGCT TTTCTCGTTG CCCATTTTAA GAATGAAGGG ATGGTCTGCG AGCATCCAGC 840
CGGGGTGGAG CCTTCTATG CTGTCTGGA AAGCCCGAGC TTCTCTCTGA TGGGCATCCT 900
40 GCTGCGGATC GCCAGTGGGA CTGCGCTCTC CATCCCATC ACTTCCAACA CATTGATCTA 960
GTCTACCCCC CACCCCGAAG ATATTAAAGT CCACTGTATC CTGTGTCCTA GGGAGCGCTT 1020
CTCTAACAAAG GCGATAGATG ATGAGGAAGA TCGTTCCAT GGTGTGCGCC TGCAGACTTC 1080
GCCCCCAATG GAACCCCTGA ACTTTGGTTC CAGTTATATT GTGTCTAATT CTGCTAACCT 1140
GAAAGTAAAG CCCAAGAGAT TGAATTTGTC CTACAGGAGC CTGGAGAAA TTCAGCACTT 1200
45 CTCAAAATTC TATGCTGGGC AGATGAAGGA ACCCATTCAA CTTCAGATTA CTGAAAAAG 1260
ACATGGGACT TTGGTGTGGG ATACTGAGGT GAAGCCAGTG GATCTCCAGC TTTAGCTGCG 1320
ATCAGCCCTC CTCTCTTTCT CAGGTGCAGC CTTTGTGAAG GAGAACCACC GGCACCTCCA 1380
AGCCAGGATG GGGAGCTGTA AAGGGGTGCT CGATGATCTC CAGGACAATG AGGTTCTTAC 1440
TGAGATGAG AAGGAGCTGG TGGAGCAGGA AAAGACACGG CAGAGCAAGA ATGAGGCCTT 1500
50 GCTGAGCATG GTGAGAGAAG AAGGGGACCT GGCCCTGGAC GTGCTCTTCA GAAGCATTAG 1560
TGAAAGGAGC CCTTACCTCG TGTCCTATCT TAGACAGCAG AATTGTGAAA ATGAGTCAGT 1620
TAGTAGTCTT GGAAGAGAGA ATCCAGCGTT CTCATTGGAA ATGGATAAAC AGAAATGTGA 1680
TCATTGATTT CAGTGTTCAG GACAGAAGAA GACTGGGTAA CATCTATCAC ACAGGCTTTC 1740
AGGACAGACT TGTAACTGG CATGTACCTA TTGACTGTAT CCTCATGCAT TTTCTCAAG 1800
55 AATGTCTGAA GAAGGTAGTA ATATTCTTTT TAAATTTTTT CCAACCATTC CTGTATATAT 1860
CACATTTTGA TCCATTGACA TGATTCTTGA AGACCCAGGA TAAAGGACAT CCGATAGGT 1920
GTGTTTATGA AGGATGGGGC CTGGAAGGAC AACTTTTCTT GATTAAATGT AAAAATAATT 1980
CCTATGGACA CTCCGTTTGA AGTATCACCT TCTCATAACT AAAAGCAGAA AAGCTAACAA 2040
AAGCTTCTCA GCTGAGGACA CTCAGGGCAT ACATGATGAC AGTCTTTTTT TTTTTGTAT 2100
60 GTTAGGACTT TAACACTTTA TCTATGGCTA CTGTTATTAG AACAAATGTA ATGTATTTCG 2160
TGAAAGAGAG CACAAAAATG GGAGAAAAATG CAAACATGAG CAGAAAAATAT TTTCCACTG 2220
GTGTGTAGCC TGCTACAAGG AGTTGTGTGG TTAATGTTC ATGGTCAACT CCAAGGAATA 2280
GTGAGATGAA ATGTGGTAAA TCAACTCCAC AGAACCAACA AAAAGAAAAAT GAGGGTAATT 2340
CAGCTTATTC TGAGACAGAC ATTCCTGGCA ATGTACCATA CAAAAATAA GCCAATCTG 2400
65 ACATTGGAT TCTACCATG ACTCTGTCA TTTGTAGCCA TTYCAGCTGT CTTTTGATTA 2460
ATGTTTCTGT GGCACACATA TTTCCATCCT TTTATGTTTA ATCTGTTTAA AACAAATGCC 2520
TAGTAGACAC CATCTGGTTG AGTCAGTTTT TTTTATGGTG TATTTTGAAC CCAATCTGAT 2580
AGTCTCTTTT AACTGGAAGA TTTCAATTAC TTAOGTTAAT GTAATTATTA ATATGTTAGG 2640
ATTTATCTCT AGTCAGCCAG TTTGTATATG TTTTCTTATT CTACTGTTAT CACATTGTGA 2700
70 CCACCTTAAAG TGGAACTTAG GCACCTTATC ACCATTTAGA TCCTATTACC TTTCTCATC 2760
TAGGATATAG TTATCTTCTA CATAATCTTT CTGTATCTTA AAACCCATCA ATAAATATT 2820
ATATATCTTT TACTTTTAAT CACTCAGAAG ATTTAAAAAA CTCATGAGAA GAGTAATCTG 2880
TTATGTTTTT CCAGATATTT ACCATTCTG TGTCTCTTCC TTTATATTAT TCCAAATTTT 2940
75 GTTCTGCAAA TTTCTCACTC TTTCTGATAGA GGTTTTTTAG TCTTTTGA GGTGTTCTGA 3000
TAGGTACAGA TTCTCTTATT TTTTGTCTCC TCTGAGGACA TCTTTTCTC ACCTTCATTC 3060
TCAGTATGAT TTTTGTCTG TAGTATTTTT AGTTGACATT GTTTTCTGTT CAGCAGTTTC 3120

5 CTTTTCAGCTT CGGTATTTC TGATGAGAAA TCTGCAGTCA TTCAAATTGT TGTTCCTCG 3180
 TATGTAGTGT GTCAATTTTC TGTCAGATTT CAAGGTATTT ATCTTTAGTT TTAGCCATT 3240
 TCATTATGTT GGGGATGAGT TTCTTGTTT TATTCCTTTT GGAATTTGCT CCAATTCATA 3300
 AATTTCAGT TTTATGTCTT TTACCAAACT TAGAGGTTTT CAGCCTAATT TCTAAAAATA 3360
 CTTTATTATA GCGTGAATTT CATCTTTATA GGAATAGTTT TAAGTATGA CAAGTTCCAA 3420
 TAGCTTATAT GCCCAGAGGG CCTTCAAAT AAGAATTTTG AAGAATACA GAAACAAAC 3480
 TTTTATATCC TTCTCATGTC TTCTACTGTA AAATTCATAT GCTTTGCTAC TCTAAACCTA 3540
 GTTTGAAATC AACAGTCTTG AGAATAGATG AAAATTTTGA TGAATAGTGG AATCTTTTA 3600
 AATGGAACCC TCTTACATGT GATTTTCCTT GCCATCTAGA AATAAACCAT AGTATTATTG 3660
 10 TTGAATCAAT CAATATTATA TTTTGTTTT TTCTCTCTCT TCTGAGACTC TTATTGTGGA 3720
 AATGTAGAGT TTTTATGTTT TCCTAAATGT CCTGATATT CTACTTATTT AGAACATCTT 3780
 TTCATTTTTT CCATTATTCT GATTGGGTAA TTTTAATTTG TCTATTTTCA AATTGTCTGG 3840
 AGTGTTCACC TGTGTGTGTC TGTGTCTGCC CACTGAGTGC ATTCAACACC TTTTAAATTT 3900
 TGTGTCAGT ATGTATCAGT TCTAAATTTT CCAATTTGTT CTCTATATTT TAAATTTCTT 3960
 15 GGCCTTATAT TTTTCTCTCT GCAATGTGT CAGCATTTGC TTGTTTGAGC TTTTCTTTT 4020
 TCAAGACAGG GTCTCAACTC TGTACCAG GCTGAGTGC AGTGTGTGCA TCTCAGCTCA 4080
 CTGCAACCTC TGCTCTCTGG TTCAAGCAT TATTGTGCT CAGCCTCTG AGTAGCTGGG 4140
 ATTAGAGCCA TGCAACACCA CAGCCAGCT AATTTTGTG ATTTTATGTA GAGACAGAGT 4200
 TTTGCTATGT TGGCCAGGCT GGTTTTGAAC TCCTGGCCTC AAGTATCCA CCCACCTCAG 4260
 20 CCTCCCAAG TCGTGGGATT ACAGGCCACT ACACCTGGCA CATTGAGTA TTTTCTTTT 4320
 TTTTCTTTT TTGAGATGGA GTCTCGCTCT GTCACTAGG CTGGAGTGA GTGTGTGAT 4380
 CTCAGCTCAC TGACGCTCT GTCTCCCGG CTCAGCGAT TCTCTTGGCT CAGCCTCTG 4440
 TAGTAGTAGG ACTACAGGTG CATGCCAACA CGCCCGGCTA ATTTTCTTAA AAAATATTTT 4500
 25 TAGTAGAGAC AGGCTTTTAC CATTTTGGCC AGGATGTTCT CGATCTCCTG ACCTCATGAT 4560
 CCACCCGCT CGGCTTCCA AAGTGTCTGG ATTACAGGCA TGAGCCACCG TGCTTGGCCT 4620
 CATTTAGATA TTTTATAAT GTCTCTTTTA AAGTCTTTG CAGATAATTC CACTGTACAT 4680
 GTTATTCAGT GTTTGTGTG CACTGAGTTC TCATTTGCCA GACAAGTGA GATTTTGTGA 4740
 GCTCATCCTT GTATCTCTAG TAGTCCGAT ATGTACCTTC GACATGTGAA TGTATCTTA 4800
 30 TGAGACTCTG TTTTATTGTT ATCCAACAGA AGATGTTTAT TATTATTTG GCTTTCTGTG 4860
 AACTGAGGTC TTAATATCAG CTCAATTTAA AAGTCTTTG AGTGTATTG GATCTATCC 4920
 TGTGTGTGCC TATGAGATG GGTGCAGTGT ATCTGTGTAG CTCCATTTCT AGGCGCTTG 4980
 AATGTGAATT AGGACACGG CAATGAATGC TCAAGTTGGG GTTGGCGTT AGAATTCATA 5040
 AAAGTCTTTA TATGCTCAG

Seq ID NO: 83 Protein sequence:
 Protein Accession #: NP_055774

40 1 11 21 31 41 51
 MMRQRQSHYC SVLFLSVNYL GGTFFGDIQS EENQIVSSYA SKVCFEIBED YKNRQFLGPE 60
 GNVDFRLIDK STNRYSVWFP TAGWYLSAT GLGLVLRDEV VTIIAFGSWS QHLALDLQHH 120
 45 EQWLVGQPLF DVTAEPEEAV ABIHLPHFIS LQGEVDVSWF LVAHPKNEGM VLSHPARVBP 180
 FYAVLESPPF SLWILLRIA SGTRLSIPIT SMTLIYHNPB PEDIKFHLYL VPSDALITKA 240
 IDDBEDRPHG VRLQTSPME PLNFGSSYIV SNSANLKVMP KELKLSYRSP BRIQHFSEKY 300
 AGQMKBPQL EITKRHGT L VMDTEVKPVD LQLVAASAPP PFGGAAPVKE NHRQLQARMG 360
 DLKGVLDLQ DNEVLTENEK BLVEQEKRTRQ SKNEALLSMV EKKGDLALDV LFRSISERDP 420
 50 YLVSYLQQN L

Seq ID NO: 84 Nucleotide sequence:
 Nucleic Acid Accession #: NM_007036
 Coding sequence: 56-610 (underlined sequences correspond to start and stop codons)

55 1 11 21 31 41 51
 CTTCCACCA GCAAGACCA CGACTGGAGA GCGAGCCGG AGGCAGCTGG GAAACATGAA 60
 GAGCGTCTTG CTGCTGACCA CGCTCTCGT GCCTGCACAC CTGCTGGCCG CCTGGAGCAA 120
 60 TAATTATGCG GTGGAAGTCC CTCAACACTG TGACAGCAGT GAGTGCAAAA CGAGCCGCGG 180
 CTGCAAGAGG ACAGTGTCTG ACGACTGTGG CTGCTGCGA GTGTGCGCTG CAGGGCGGGG 240
 AGAAACTTGC TACGACACAG TCTCAGGCAT GGATGGCATG AAGTGTGGCC GGGGCTGAG 300
 GTGTCAAGCT TCTAATGGGG AGGATCCTTT TGGTGAAGAG TTTGGTATCT GCAAAGACTG 360
 65 TCCTCAGCGC ACCTTCGGGA TGATTCAGAG AGAGACCTGC AACTGCCAGT CAGGCATCTG 420
 TGACAGGGGG ACAGGAAAT GCCTGAAAT CCCCTCTCTC CAATATTCTG TAACCAAGTC 480
 TTCCAACAGA TTTGTTTCTC TCAAGGAGCA TGACATGGCA TCTGGAGATG GCAATATTGT 540
 GAGAGAAGAA GTTGTGAAG AGAATGTCTC CGGGCTCTCC GTAATGAGGA AATGTTTAA 600
 TCCACGCTGA TCCCGGCTGT GATTTCTGAG AGAAGGCTCT ATTTTCTGTA TTGTTCAACA 660
 70 CACAGCCAAC ATTTTAGGAA CTCTCTAGAT ATAGCATAAG TACATGTAAT TTTGAAGAT 720
 CCAAAATTGT ATGCATGGTG GATCCAGAAA ACRAAAAGTA GGATACTTAC AATCCATAAC 780
 ATCCATATGA CTGAACACTT GTATGTGTTT GTTAAATATT CGAATGCATG TAGATTGTGT 840
 AAATGTGTGT TATAGTAAC ACTGAAGAAC TAAAAATGCA ATTTAGGTAA TCTTACATGG 900
 AGACAGGTCA ACCAAGAGG GAGCTAGGCA AAGCTGAAGA CCGCAGTGA TCAATTAGT 960
 75 TCTTTGACTT TGATGTACAT TAATGTTGGG ATATGGAATG AAGACTTAAG AGCAGGAGAA 1020
 TGATGGGAGG GGGTGGGAGT GGGAAATAAA ATATTAGACC CTTCCTTGTG AGGTAGCTTC 1080
 TCTAGAAATT AATTGTGCTT TTTTCTTTT TTTGGCTTTG GAAAAAGTCA AAATAAAACA 1140

ACCAGAAAAC CCCTGAAGGA AGTAAGATGT TTGAAGCTTA TGGAAATTTG AGTAACAAAC 1200
 AGCTTTGAAC TGAGAGCAAT TTCAAAGGC TGCTGATGTA GTTCCCGGT TACCTGTATC 1260
 TGAAGGACGG TTCTGGGGCA TAGGAAACAC ATACACTTCC ATAAATAGCT TTAACGTATG 1320
 CCACCTCAGA GATAAATCTA AGAAGTATTT TACCCACTGG TGGTTTGTGT GTGTATGAAG 1380
 5 GTAAATATTT ATATATTTTT ATAAATAAAT GTGTTAGTGC AAGTCATCTT CCCTACCCAT 1440
 ATTTATCATC CTCTTGAGGA AAGAAATCTA GTATTATTG TTGAAAATGG TTAGAATAAA 1500
 AACCTATGAC TCTATAAGGT TTCAAACAT CTGAGGCATG ATAAATTTAT TATCCATAAT 1560
 TATAGGATC ACTCTGGATT TCMAAAATG TCMAAAATG AGCAACAGAG GGACCTTATT 1620
 TAAACATAAG TGCCTGTAAT TCAATTTTAA GGTATGAAAA TAAGTTTTTA 1680
 10 SGAGGTTTGT AAAAGAAGAA TCAATTTTCA GCAGAAAACA TGCTCACTTT AAAATATAGG 1740
 TGGAAATAGG AGTATATTTG AAAGAATCTT AGCACAACA GGACTGTTGT ACTAGATGTT 1800
 CTTAGGAAAT ATCTCAGAAG TATTTTATTT GAAGTGAAGA ACTTATTTAA GAATTTATTT 1860
 AGTATTATCC TGTATTTTAT TCTTGAAGTT GGCCAACAGA GTTGTGAATG TGTGTGGAAG 1920
 15 GCCTTTGAAT GTAAAGCTGC ATAAAGCTGT AGGTTTGTGT TTAAGAGGAC ATGTTTATTA 1980
 TTGTTCAATA AAAAGAACA AGATAC

Seq ID NO: 85 Protein sequence:

Protein Accession #: NP_008967.1

20 1 11 21 31 41 51
 | | | | |
 MKSVLLITL LVPALHVAAM SNRYAVDCPQ HCDSSBECKSS PRCKRTVLD CGCCRVCAAG 60
 RGETCYRTVS GMDGMKCGPG LRCQPSNGSD PFGEEPGICK DCFYGTGMD CRSTNCQSG 120
 25 ICDRTGKCL KPFPFQYSVT KSSNRFVSLT EHDMAAGDGN IVREEVVKEN AAGSPVMRW 180
 LNPR

Seq ID NO: 86 Nucleotide sequence:

Nucleic Acid Accession #: D86983

Coding sequence: 52-4491 (underlined sequences correspond to start and stop codons)

30 1 11 21 31 41 51
 | | | | |
 35 AGCCGCGCGT GGTGGCTCCG TCGTCCGAG CGTCCGTCG CGCGTCGGC CATGCCAAG 60
 CGCTCCAGGG GCGCCGGGGG CCGCTGCCCTG TTGGCGCTCG TGCTGTTCTG CGCTGGGGG 120
 ACGCTGGCCG TGGTGGCCCA GAAGCGGGG CGAGGGTGTG CGAGCCGCTG CCGTGTCTTC 180
 CGCACCAACG TGGCTGTCAT GCATCTGCTG CTGAGGCGCG TGCCCGCGCT GGCGCCGCG 240
 40 ACCTCCATCC TAGATCTTCG CTTTAACAGA ATCAGAGAGA TCACACTCG GGCATTCAG 300
 CGGCTGAGGA ACTTGAACAC ATTGCTTCTC AATAATATC AGATCAAGAG GATACCTAGT 360
 GGAGCATTTG AAGACTTGA AATTTTAAA TATCTCTATC TGTACAGAA TGAGATCCAG 420
 TCAATTGACA GGCAAGCATT TAAGGGACTT GCCTCTCTAG AGCACTATA CCGTCACTTT 480
 AATCAGATAG AAATTTTGA CCCAGATTCT TTCCAGCATC TCCCGAAGCT CGAGAGGCTA 540
 TTTTGTGATA ACAACCGGAT TACACATTTA GTTCCAGGGA CATTTAATCA CTTGGAATCT 600
 45 ATGAAGAGAT TGGACTGGA CTCAAACACA CTTCACCTGC ACTGTGAAT CCGTGTGTTG 660
 CGCGATTTCG TGAACAACCTA CGCGAGTGG GGAACGCGCG AGGCAGCGCG CATCTGTGAA 720
 TATCCAGAGC GCATCCAGGG ACGCTCAGTG GCAACCATCA CCGCGAAGA GCTGAACGTG 780
 GAAAGGCCCC GGATCACCTC CGAGCCCCAG GACGCAGATG TGACCTCGGG GAACACCGTG 840
 TACTTCACTC GCAGAGCCGA AGGCAACCCC AAGCCTGAGA TCATCTGGCT GCGAACAAT 900
 50 AATGAGCTGA GCATGAAGAC AGATTCCCGC CTAACTTGC TGGACGATGG GACCTGATG 960
 ATCCAGAAAC CACAGAGAC AGACCAAGGT ATCTACAGT GCATGCCAA GAACGTGGCC 1020
 GGAGAGGTGA AGACGCAAGA GGTGACCTC AGGTACTTGG GTCTCCAGC TCGACCCACT 1080
 TTTGTAATCC AGCCACAGAA TACAGAGGTG CTGGTTGGGG AGAGCGTCAC GCTGGAGTGC 1140
 AGCGCCACAG GCCACCCCCC GCGCGGATC TCGTGGACGA GAGGTGACCG CACACCTTG 1200
 55 CCAGTTGACC CGCGGGTGAA CATCAAGCCT TCTGGCGGGC TTTACATACA GAACGTGTA 1260
 CAGGGGACA GCGGAGAGTA TGGCTGCTCT GCGACCAACA ACATTGACAG CGTCCATGCC 1320
 ACCGCTTTCA TCATCGTCCA GGCTCTTCTT CAGTTCACTG TGACGCTCA GGCAGAGTC 1380
 GTTATTGAGG GCGAGACCGT GGATTTCCAG TGTGAAGCCA AGGCAACCC GCGCGCGTG 1440
 ATCGCTTGA CCAAGGGAGG GAGCCAGCTC TCGTGGACC GCGGCACTT GGTCTGTCA 1500
 60 TCGGGAACAC TTAGAATCTC TGGTGTGACC CTCCAGGACC AGGCGCAGTA GGAATGCCAG 1560
 GCTGTCAACA TCATCGGCTC CCAGAAGGTC GTGGCCACC TGACTGTGCA GCCAGAGTC 1620
 ACCCCAGTGT TTCCAGCAT TCCAGGAC ACACAGTGG AGTGGGCGC CAATGTGAG 1680
 CTCCTGTGCA GCTCCAGGG CGAGCCCGAG CCAGCCATCA CCTGGAACAA GGATGGGGT 1740
 CAGGTGACAG AAAGTGAATA ATTTACATC AGCCCTGAAG GATTCTTGAC CATCAATGAC 1800
 65 GTTGGCCCTG CAGACGCAGG TCGCTATGAG TGTGTGGCCC GGAACACCAT TGGGTGCGCC 1860
 TCGTGTGACA TGTGACTCAG TGTGAACGTT CCTGAGTCA GTCGAAATGG AGATCGTGT 1920
 GTAGCTACCT CCATCGTGA AGCGATTGGG ACTGTTGACA GAGCTATAAA CTCACCCGA 1980
 ACACATTGTT TTGACAGCG TCCTCGTTCT CCAATGATT TGCTGGCTT GTTCCGTAT 2040
 70 CCGAGGATC CTTACACAGT TGAACAGGCA CGGCGGGAG AAATCTTTGA ACGGACATTG 2100
 CAGCTCATTC AGGAGCATGT ACAGCATGGC TTGATGGTGG ACCTCAACGG AACAGTTAC 2160
 CACTACACCG ACCTGGTGTG TCCACAGTAC CTGAACCTCA TCGCAACCT GTGCGGCTGT 2220
 ACCGCCACC GCGCGTGAA CACTGCTCG GACATGTGCT TCCACGAA GTACCGAGG 2280
 CACGAGCGCA CCGTGAACAA CCTGACGAC CCCATGTGG GCGCTCGCT GACCGCTTC 2340
 75 GAGCGCTGCT TGAATCCGT GTACAGAAAT GGCTTCAACA CCGCTCGGG CATCAACCC 2400
 CACCGACTGT ACACGGGCA CGCCTTCCC ATGCGCGCC TGGTGTCCAC CACCTGTATC 2460
 GGGACGGAGA CCGTCAACCC CGACGAGCAG TTCACCCACA TGCTGATGCA GTGGGGCCAG 2520

	TTCTCTGACC	ACGACCTCGA	CTCCACGGTG	GTGGCCCTGA	GCCAGGCAGC	CTTCTCCGAC	2580
	GGACAGCACT	GCAGCAACCT	GTGCAGCAAC	GACCCCCCTT	GCTTCTCTGT	CATGATCCCC	2640
	CCCAATGACT	CCCGGGCCAG	GAGCGGGGCC	CGCTGCATGT	TCTTCTGTGT	CTCCAGCCCT	2700
5	GTGTGCGGCA	GCGGCATGAC	TTGCTGTCTC	ATGAACCTCG	TGTACCCGCG	GGAGCAGATC	2760
	AACCAAGCTCA	CTCTCTACAT	CGACGATCCC	AACGTGTACG	GGAGCAGCGA	GCATGAGGCC	2820
	CGCAGCATCC	GCGACCTGGC	CAGCCACCGC	GGCTGTCTGC	GGCAGGGCAT	CGTGACAGCG	2880
	TCCGGGAAGC	CGCTGCTCCC	CTTCCGCCAC	GGGCCGCCCA	CGAGTGTGAT	GCGGGACGAG	2940
	AACGAGAGCC	CCATCCCTTG	CTTCTTGGCC	GGGACCAACC	GCGCCAACGA	GCAGCTGGGC	3000
	CTGACCAAGCA	TGCACAGSCT	GTGGTTCCGC	GAGCACAACC	GCATTGCCAC	GGAGCTGCTC	3060
10	AAGCTGAACC	CGCATGGGA	CGCGACACCC	ATCTACTATG	AGACCAAGAA	GATCGTGGGT	3120
	GCGGAGATCC	AGCAGATCAC	CTACCAGCAC	TGGCTCCCGA	AGATCCTGGG	GGAGGTGGGC	3180
	ATGAGGACGC	TGGAGAGGTA	CCAAGGCTAC	GACCCCGGCA	TCAATGCTGG	CATCTTCAAC	3240
	GCTTTCGCCA	CCCGGCCCTT	CAGGTTTGGC	CACACGCTTG	TCAACCCACT	GCTTTACCGG	3300
	CTGACAGAGA	ACTTCCAGCC	CATTGCACAA	GATCACTCTC	CCCTTCAACA	AGCTTTCTTC	3360
15	TCCTCCCTTCC	GGATTGTGAA	TGAGGGGSGC	ATCGATCCGC	TTCTCAGGGG	GCTGTTCCGG	3420
	TGCGCGGGGA	AAATGCGTGT	GGCTTCGCAG	CTGCTGAACA	CGAGCTCAC	GGAGCGGCTG	3480
	TTCTCCATAG	CACACAGGTT	GGCTCTGGAC	CTGGCGGCCA	TCAACATCCA	GCGGGGCGGG	3540
	GACCAAGGGA	TCCCACTCTA	CCAGACTTAC	AGGGTCTACT	GCAATCTATC	GCGGCGCACAC	3600
20	ACGTTGAGGG	ACCTGAAAAA	TGAGATTAAA	AACCTTGAGA	TCCGGGAGAA	ACTGAAAAAG	3660
	TTGTATGGCT	CGACACTCAA	CATCGACCTG	TTTCCGGGCG	TGTTGGTGA	GGAGCTGGTG	3720
	CTTGGCAGCC	GGCTGGGCC	CACCTGATG	TGTCTTCTCA	GCACACAGTT	CNAGCGGCTG	3780
	CGAGATGGGG	ACAGGTTGTG	GTATGAGAAC	CGTGGGTTGT	TCTCCCGCGC	CCAGCTGACT	3840
	CAGATCAAGC	AGAGTCTGCT	GGCCAGGATC	CTATGGGACA	ACGCGGACAA	CATCACCCGG	3900
25	GCTTCAGAGG	ACGTTGTTCAG	GGTGGCGGAG	TTCCCTCAGG	GCTACGGCAG	CTGTGACGAG	3960
	ATGCCCAAGG	TGAGCTCCG	GGTGTGGCAG	GACTGTCTGT	AAGACTGTAG	GACCAAGGGG	4020
	CAGTTCAATG	CTTTTCCCTA	TCAATTTCCGA	GGCAGACGGT	CTCTTGAATT	CAGCTACCA	4080
	GAGACCAAGC	CGACCAAGAA	AACAAGACCA	CGGAAATATC	CCAGTGTGAG	GAGACAGGGG	4140
	GAAATCTCCA	GCACACGAC	CTCAGCTTTC	AGCACACGCT	CAGATGCATC	TGGGACAAAT	4200
30	GACTTCAGAG	AGTTTGTCT	GGAAATGCAG	AAGACCATCA	CAGACCTCAG	AACACAGATA	4260
	AAGAACTCTG	AATCAAGGCT	CAGTACCACA	GAGTGGTGG	ATGCGGGGG	CGAATCTCAC	4320
	GCCAACTACA	CCAGTGGGAA	AAAAGATGCA	TGCACCATTT	GTGAATGCAA	AGACGGGCG	4380
	GTCACCTGCT	TGCTGGAAGC	TTGCCCCCTT	GCCACCTGTG	CTGTCCCGCT	GAACATCCCA	4440
	GGGCGCTGCT	GTCCAGTCTG	CTTACAGAA	AGGGCGGAGG	AAAAGCCCTA	GGCTCTCGGG	4500
35	AGGCTCTCTA	GAGTTTGTCT	GCTGTGCCAT	CGTGAGATCG	GGTGGCCGAT	GGCAGGAGGC	4560
	TCCGAGCTGC	AGACCAAGAA	ACACCCAGAA	CTCGTGACAT	TTCTATGCAA	CGTCCAGCTG	4620
	GTGCTGTATC	AGAAGGCAGT	CGAGGAGGCT	TCCAACAGCA	GCATCTGGGG	AGAAGGAGGC	4680
	ACAGCAGGTT	CCTGAAGGGA	AGCAGGCGAG	AGTCTAGCT	TCACTTAGA	CTTCTCAGGT	4740
	TTTATTTTAA	TTCTTTTAAA	ATGAAAAAAT	GGTGCTACTA	TAAATTTGCA	CAGTTGAATC	4800
40	TTTATAGGCG	CTAAATGGT	TTTGCCCTCC	AACACCATTT	CTTTTAAAT	AAAGCAGGAT	4860
	AGCTCTATAT	GTGAGCCTTG	CCTTGTTCAG	ATGCCAGGAG	CGCGCAGACC	TGTACCCCGC	4920
	QGGGCGGCTG	AGTCTCGGAG	CTGCCAGAGG	GGCTCACCGA	AATCGGGGTT	CCATCACAA	4980
	CTATGTTTAA	AAAGAAAAAT	GGTGTTTGGC	AAACGGAACA	GAACCTTTGA	TGAGAGCGTT	5040
	CACAGGAGCA	CTGTCTGGGG	GTGAGTGCA	AGCCCCCGGC	CTCTTCCCTG	GGAACTCTG	5100
45	AACTCTCTCT	TCCTCTGGGC	TCTCTGTAA	ATTTACCCAC	ACGTGAGCAT	CTAATCCCAA	5160
	GACAAACATT	CCCGCTGCTC	GAGCAGCTG	TATAGCCTGT	GACTCTCCGT	GTGTGACGTC	5220
	CTTCCACACC	TGATTAGAAC	ATTCATAAGC	CACATTTAGA	AACAGATTTG	CTTTTCACTG	5280
	TCACTTGAC	ACATACTGCC	TAGTTGTGAA	CCAAATGTGA	AAAAACCTCC	TTCATCCCAT	5340
	TGTGTATCTG	ATACCTGGCG	AGGGCCAAAG	GTGTGTGTGT	ACAAAGCCCG	TCCAGCCCGG	5400
50	CCCTGTGCTG	GTCCACGTCC	TGAACAAGAG	CCGCTTCCGG	ATGGCTCTTC	CCAAGGGAGG	5460
	AGGAGCTCAA	GTGTGGGAA	CTGTCTAACT	TCAGGTTGTG	TGAGTGCCTT		

Seq ID NO: 87 Protein sequence:

Protein Accession #: BAAL3219

	1	11	21	31	41	51	
55	SRPFWLRASE	RPSAPSAMAK	RSRPGRRRL	LALVLPFANG	TLAVVAQKPG	AGCPSRCLCF	60
	RTTVRCMHLL	LEAVPAVAPQ	TSILDLRPNR	IREIQPGAFR	RLRLNLTLL	NNNQIKRIPS	120
	GAFEDLENLK	YLVLKNEIQ	SIDRQAFKGL	ASLEQLYLHF	NQISTLDPS	FOHLPKLERL	180
60	FLHNNRITHL	VPOTFNHLES	MKRLRLDSNT	LHDCCELNL	ADLLKTYAES	GNAQAAACE	240
	YPRRIQGRSV	ATITPEELNC	ERPRITSEPO	DADVTSGNTV	YFTCRAGNPN	KPEIILNLRN	300
	NELSMKTDNR	LNLLDDGTLM	IQNTQSTDOG	IYQCMKQNV	GEVKTQEVTL	RYFGSPARPT	360
	FVIQPMTEV	LVGESVTLEC	SATGHPPPRI	SWTRGRDPL	PVDPRVNIPT	SGGLYIQNVV	420
	QDGSGEVACS	ATNNDISVHA	TAPIIVQALP	QFTVTPQDRV	VIEGGTVDPQ	CEAKGNPPPV	480
	IATWKGQSQL	SVDRHLVLS	SGTLRISGVA	LHDQQQYECQ	AVNIGSQKV	VAHLTVQPRV	540
65	TPVPASIPSD	TTVEVGANVQ	LPSCSQGEPE	PAITWNKDG	QVTESGKPHI	SPBGFITIND	600
	VGPADAGRYE	CVARNTIGSA	SVSMVLSENV	PDVSRNGDPF	VATSIIVEAIA	TVDRAINSTR	660
	THLPDSRPRS	PNDLLALFRY	PRDPYTVBQA	RAGRIPERTL	QLIQEHVQHG	LMVDLNGTSY	720
	HYNDLVSPQY	LNLIANLSCG	TAHRRVNNCS	DMCFHQKYRT	HDGTCCNLQH	PMMGASLTAF	780
	ERLLKSVYEN	GPNTPRGINP	HRLYNGHALP	MPLVSTTLI	GTETVTPDEQ	PTHMLMQWQ	840
70	FLDHDLSILL	VALSQARFSD	GQHCNSVCSN	DPCCFSVMIP	PNDSTRARSGA	RCMPFVRSSP	900
	VCGSGMTSLL	MNSVYPREI	NQLTSYIDAS	NVYGSTHEEA	RSIRDLSHR	GLLRQGIQVR	960
	SKPLPLFPAT	GPPTCMRDE	NESPIPCFLA	GDHRANEQLG	LTSMTLWFR	EHRIATELL	1020
	KLNPHMDQGT	IYYETRKIVG	AEIQHITVQH	WLPKILGEVG	MRTLGEYHYG	DPGINAGIFN	1080
	AFATAAPRPG	HTLVNPLLYR	LDENFQPIAQ	DHLPLHKAFF	SPFRIVNEGG	IDPLRLGLFG	1140
75	VAGKMRVPSQ	LNATELTERL	FMAHTVALD	LAANIQGRG	DHGIPPYHGY	RVYCNLSAAH	1200
	TFEDLNKNIK	NPEIREKLKR	LYGSTLNIDL	FPALVVEDLV	PGSRIGPTLM	CLLSQVFKRL	1260

RDGRLMYEN PGVPSPAQLT QIKQTSIARI LCDNADNITR VQSDVPRVAS FPHGYGSCDE 1320
 IPFVLDLRVWQ DCCEDCKTRG QFNAPSYHFR GRRSLFYSYQ SDKPTKTRP RKIPSVGRQG 1380
 EHLNSNSTAF STRSDASGTN DFRBFLVEMQ KTTDLRTQI KKLBSRLSTT ECVDAAGBSH 1440
 ANNTKWKIDA CTICBCKDQ VTCFVRACPF ATCAVPVNI P GACCPVCLQK RAREKP

Seq ID NO: 88 DNA sequence

Nucleic Acid Accession #: NM_004834.1

Coding sequence: 80-3577 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 15 AATTCGAGGA TCCTGGTACC ATGGCACAGA GCGACAGAGA CATTTATTGT TATTGTGTTT 60
 TTGGTGGCAA AAAGGGAAA TGCGGAACGA CTCCCTGCA AAAAGTCTGG TGACATCGA 120
 CCTCTCTCC TCCTGGGATC CTCTGGGATG TTTTGAAGCTG GTGGAAGTGG TTGGAATGG 180
 CACTATGGA CAAGTCTATA AGGTCGACA TGTAAAAAG GGTCAATGG CAGCCATCA 240
 AGTATGAGAT GTCACTGAGG ATGAAGAGGA AGAAATCAAA CTGGAGATAA ATATGCTAAA 300
 GAAATACCTC CATCACAGA ACATTGCAAC ATATTATGCT GCTTTCATCA AAAAGAGCCC 360
 20 TCCAGGACAT GATGACCAAC TCTGGCTTGT TATGGAGTTC TGTGGGGCTG GGTCCATTAC 420
 AGACCTTGTG AAGAACAACA AAGGGAACAC ACTCAAGAA GACTGGATCG CTACATCTC 480
 CAGAGAAATC CTGAGGGGAC TGGCACATCT TCACATTCAT CATGTGATTC ACCGGGATAT 540
 CAAGGGCCAG AATGTGTGTC TGACTGAGAA TGCAGAGGTG AAACCTGTGT ACTTTGGTGT 600
 GAGTGTCTAG CTGGACAGGA CTGTGGGGCG GAGAAATACG TTCATAGGCA CTCCTACTG 660
 25 GATGGCTCTC GAGGTCACTG CCTGTGATGA GAACCCAGAT GCCACCTATG ATTACAGAAG 720
 TGATCTTTGG TCTGTGGCA TTACAGCCAT TGAGATGGCA GAAGGTGCTC CCCCTCTCTG 780
 TGACATGCAAT CCAATGAGAG CACTGTTTCT CATTCACAGA AACCTCTCTC CCGGCTGAA 840
 GTCAAAAAA TGTCGAAGA AGTTTITTAG TTTTATAGAA GGGTGCCTGG TGAAGAATTA 900
 CATCAGCGGG CCTCTACAG AGCAGCTTTT GAAACATCTT TTTATAGGG ATCAGCCAAA 960
 30 TGAAAGGCAA GTTAGAATCC AGCTTAAGGA TCATATAGAT CGTACCAGGA AGAAGAGAGG 1020
 CGAAGAAATG GAACTGAGT ATGATACAG TGGAGTGAAG GAAGAAGAGG AGAAGTGGCC 1080
 TGAACAGGAA GGAGAGCCAA GTTCCATTGT GAACTGTCTT GGTGAGTCTA CTCTTCGCGG 1140
 AGATTTCCTG AGACTGCAGC AGGAGAACAA GGAACGTTCG GAGGCTCTTC GGAGACAA 1200
 GTTACTACAG GAGCAACAGC TCCGGGAGCA GGAAGAATAT AAAAGGCAAC TGTGTCGAGA 1260
 35 GAGACAGAA GCGATTGAGC AGCAGAAAGA ACAGAGGCGA CCGCTAGAAG AGCAACAAAG 1320
 GAGAGAGCGG GAGGCTAGAA GGCAGCAGGA ACCTGAAACG CGAAGGAGAG AACAAAGAA 1380
 AAAGAGGCGT CTAGAGAGT TGGAGAGAA GCGCAAGAA GAAGAGGAGA GGAAGCGGGC 1440
 AGAAGAGAGG AAGAGGAGAG TTGAAGAGA ACAGGAGTAT ATCAGGCGAG AGCTAGAGAA 1500
 40 GAGAGAGAGG CACTTGAAG TCCTTCAGCA GCAGCTGCTC CAGGAGCAGG CCACTGTACT 1560
 GCATGACCAT AGGAGGCGCG ACCCGCAGCA CTGCGCAGC CCGCCACCAC CGCAGCAGGA 1620
 AAGAGAGGAA CCAAGCTTCC ATGCTCCCGA GCCCAAGGCC CACTACGAGC CTGCTGACCG 1680
 AGCGCGAGAG GTTCTCTGTA GAACAACATC TCGCTCCCTT GTTCTGTCCC GTGAGATTC 1740
 CCCACTGCGG GGCAGTGGGC AGCAGAAATG CCAAGCAGGA CAGAGAAACT CCACCATAT 1800
 45 TGAGCCGAGG TTTCTGTGG AGAGAGTGGG GAAGCTGGTG CCCAGACCTG GCAGTGGCAG 1860
 CTCCTCAGGG TCACGCAACT CAGGATCCCA GCCCGGTCTC CACCTGGGTG CTCAGAGTGG 1920
 CTCGCGGAGG CCGTTCAGAG TGAGATCATC ATCCAACTCT GAAGGCTCTC CATCTCAGCG 1980
 CCTGGAAAT GCAGTGAAGA AACCTGAAGA TAAAGAGGAA GTTTTCAGAG CCTCTAAGCC 2040
 TGCTGGCGAA GTGGATCTGA CCGCACTGGC CAAAGAGCTT CGAGCAGTGG AAGATGTAG 2100
 50 GCCACTCTAC AAGTAACGG ACTACTCTCT ATCCAGTGA GAGTCCGGGA CGACGGATGA 2160
 GGAGGACGAC GATGTGGAGC AGGAAGGGGC TGACAGTCC ACCTCAGGAC CAGAGGACAC 2220
 CAGAGCAGCG TCATCTCTGA ATTTGAGCAA TGTGAAAGC GAACTCTGTA AAACCATGAT 2280
 TGTCCATGAT GATGTAGAAA GTGAGCGCGC CATGACCCCA TCCAAGGAGG GCACTCTAAT 2340
 CGTCCGCGAG ACTCAGTCCG CTAGTAGCAC ACTCCAGAAA CACAAATCTT CCTCTCTCTT 2400
 55 TACACCTTTT ATAGACCCCA GATTACTACA GATTCTCCA TCTAGCGGAA CAACAGTGAC 2460
 ATCTGTGGTG GGATTTTCCT GTGATGGGAT GAGACCAGAA GCCATAAGGC AAGATCTCTAC 2520
 CCGGAAGAGC TCAGTGGTCA ATGTGAATCC TACCACACT AGGCCACAGA GTGACACCCC 2580
 GGAGATTCTT AAATACAGA AGAGGTTTAA CTCTGAGATT CTGTGTGCTG CCTTATGGGG 2640
 AGTGAATTTG CTAGTGGGTA CAGAGAGTGG CCTGATGCTG CTGGACAGAA GTGGCCAAAG 2700
 GAAGGTCTAT CCTCTTATCA ACCGAAGACG ATTTCAACAA ATGGACGTAC TTGAGGGCTT 2760
 60 GAATGTCTTG GTGACAAATC CTGGCAAAAA GGATAAGTTA CGTGTCTACT ATTGTCTCTG 2820
 GTTAAAGAAAT AAAATACTTC ACAATGATCC AGAAGTTGAG AAGAAGCAGG GATGGACAAC 2880
 CGTAGGGGAT TTGGAAGGAT GTGTACATTA TAAAGTTGTA AAATATGAAA GAATCAAAAT 2940
 GCTGTGATTT GCTTTGAAGA GTTCTGTGGA AGTCTATCG TGGGCACCAA AGCCATATCA 3000
 65 CAATTTATG GCTTTAAGT CATTTGGAGA ATTGATACAT AAGCCATTAC TGGTGGATCT 3060
 CACTGTGTAG GAAGCCAGA GGTGAAAGT GATCTATGGA TCCTGTGCTG GATTCATATG 3120
 TGTGTATGTG GATTCAAGAT CAGTCTATGA CATTTATCTA CCAACACATG TAAGAAAGAA 3180
 CCCACACTCT ATGATCCAGT GTAGCATCAA ACCCATGCA ATCATCATCC TCCCAATATC 3240
 AGATGGAATG GAGCTCTGG TGTCTATGA AGATGAGGGG GTTTATGTAA ACACATATGG 3300
 70 AAGGATCACC AAGGATGTAG TTCTACAGTG GGGAGAGATG CCTACATCAG TAGCATATAT 3360
 TCGATCCAAT CAGACAATGG GCTGGGAGA GAAGGCCATA GAGATCCGAT CTGTGGAAC 3420
 TGGTCACTTG GATGTGTGT TCATGCAAAA AAGGGCTCAA AGACTAAAAT TCTGTGTGTA 3480
 ACGCAATGAC AAGGTGTCTT TTGCTCTGTG TCGSTCTGGT GGCAGCAGTC AGGTTTATTT 3540
 CATGACCTTA GGCAGGACTT CTCTTCTGAG CTGATGAGAG CAGTGTGATC CAGGATATAC 3600
 75 TGGCCTCCAG AGTCTTCAAG ATCCTGAGAA CTGGAATTC CTGTAACTG GAGCTCGGAG 3660
 CTGCAACGAG GGCACCAAG ACAGCTGTGT GTGCAGACCT CATGTGTGG GTTCTCTCCC 3720
 CTCCTCTCTG TTCTCTTAT ATACAGATT ATCCCCATTC TTTTITTTTT TCTTACTCCA 3780

AAATAAATCA AGGCTGCAAT GCAGCTGGTG CTGTCAGAT TCCAAAAA AAAAAAACC 3840
ATGGTACCG GATCTCGAA TTCC

5 Seq ID No: 89 Protein sequence:
Protein Accession #: NP_004825.1

```

10 1      11      21      31      41      51
    |      |      |      |      |      |
MNDSPAKSL VDDLSSLRD PAGIFRLVEV VGNQTYCQVY KGRHVKTQQL AAIKVMQVTE 60
DEBBIKLEI NMLKKYSHHR NIATYYGAFI KKSPPGHDDQ LNLVMSFCGA GSITDLVKNT 120
KGNLTKEWI AYISREILRG LAHLHIHVI HRDIKGQNLV LTERAEVKLV DFGVSAQLDR 180
TVGRNFTFIG TPYWMAPFVI ACDEPNPATY DYRSDLWSCG ITAIRMABGA PFLCDMHFMR 240
ALFLIPRNPP PRLKSKKMSK KFFSFIEGCL VKNYMQRPST BQLKHPFIR DQPHRQVRI 300
QLKDHIDRTR KKRGEKDETE YEYSGSEBEE BEVPEQEGEP SSIVNVPGES TLRRDFLRLO 360
QNKERSEAL RRQQLQEQQ LREQSEYKRO LLAERQKRIE QQKQRRLLE EQQRREAR 420
RQQRERQRRR EQREKRLLE LERRRKEBEE RRAERKRRR VEREQEYIRR QLEBQRHLE 480
VLQQLLQEQQ AMLLHDHRRP HPQHSQQPPP PQERSEKPSF HADEPKAHYE PADRAREVTV 540
RTTSRSPVLIS RRDSPLQSGS QONSQAGQPN STSIEPRILW ERVEKLVPRP GSGSGSGSSN 600
SQSQPGSHPG SQSGSGSERFR VRSSSKSEBS PSQRLNNAV KPEDKKEVPR PLKPAGEVDL 660
TALAKELRAV EDVRPPHKVT DYSSSSSESG TDBEDDDVE QSGADESTSG FEDTRAASSL 720
NLSNGTESVS KTWIVHDVSE SEPAMTPSKE GTLIVRQTS ASSTLQKHKS SSSFTPPIDP 780
RLGLISPSGS TTVTSVVGFS CDGMRPRAIR QDPTRKGSVV NVNPTNRPQ SDTPBIRKYK 840
KRFNSILCA ALWGVNLLVG TESQLMLLDR SQGKVPPLI NRRRPQQMDV LEGLIVLVTI 900
SKKDKLRVY YLSWLRNKIL HNDPEVKKQ GWITVGDLEQ CVHYKVKVYE RIKFLVIALK 960
SSVEVYAMAP KPYHKMAFK SFGLVHKPL LVDLTVESQ RLKVIYGSQA GFHAVDVDSG 1020
SYDIYLPHT VRKNSHSMIQ CSIKPHALII LFNIDGMELL VCYEDGVYV NTYGRITKDV 1080
VLQNGRMPIS VAYIRSNQTM GWGEKAIEIR SVETGHLGV PMHKRAQRLK FLCERNKVP 1140
30 FASVRSKSSS QVYPTLGLRT SLLSW

```

Seq ID NO: 90 DNA sequence
Nucleic Acid Accession #: none found
35 Coding sequence: 2-71 (underlined sequences correspond to start and stop codons)

```

40 1      11      21      31      41      51
    |      |      |      |      |      |
TTACTCTTCA ATTCTTACA CGGTATTCA AACAAACAGT TTGCTGAGA GGAGCTTTTG 60
TCTCTCTCTA AGAAAAGTTT TATAAAGCTG AAAGGAATC AACACGTAAT CTTAAAAATG 120
AAAACAAAC AAACCAACAA CCTAGATAAC TACAGTATC AGGAGACACA GTTCAACTCC 180
TTGTTATGTT TTAGTCATAT GGCCTACTCA AACAGCTAAA TAACAACACC AGTGGCAGAT 240
AAAAATCACC ATTATCTTT CAGCTATTAA TCTTTGAAT GAATAAATG TGACAAACAA 300
45 ATTAACATT TTGAACATGA AAGGCAACT CTGCACAATC CTGTATCCAA GCAAACTTTA 360
AATTATCCAC TTAATTATTA CTTAATCTTA AAAAAAATTA GAACCCAGAA CTTTTCATG 420
AAGCATTGGA AAGTTGAAGT GGAATTTAGG AAAGCCATAA AATATATAAT ACTGTTATCA 480
CAGCACCAGC AAGCCATAAT CTTTATACCT ATCAGTTCTA TTCTATTAA CAGTAAAAAC 540
ATTAAGCAAG ATATAAGACT ACCTGCCCAA GAATTCAGTC TTTTTCATT TTTGTTTTTC 600
TCAGTCTGGA GGATGTTAAT CGTCAAATTT TCTTTGACT GCATTCTCA CTACTTTTTG 660
CACAATGGTC TCACTTCTC ACATTTGTTT TCGGAATAA ATTGATAAAA GGTGTTAAGT 720
TCTGTGAATG TCTTTTAAAT TATGGGCATA ATTGTGCTTG ACTGATAAAA AACTTAAGTC 780
CACCTCTATG TTTATAATAA TTTCTTGAGA ACAGCAAACT GCATTATCCA TCGTAAAACA 840
ACATCTGACT TAGGGGAGCT GCAGGGAAGT GGTGAGACAG TTGGAACGGC TCCTCAGAAA 900
55 TCCAGTGACC CAATTCATAA GACCATAGCA CCTGCAAGTG ACACAACAG CAGATTATT 960
ATACATTAT TAGCCTTAGC AGGCAATAAA CCAAGATCA CTTTGAAGAC ACAGCAAAA 1020
GTGATACCT CGCAGATCT GAAATAGATG TGTCTCAGA CAACAAGTC CCTTCAGAT 1080
CTTCATGTTG CATAAATGTT ATGAATATTA ATAAAAAGTT GATTGAGA

```

60 Seq ID No: 91 Protein sequence:
Protein Accession #: none found

```

65 1      11      21      31      41      51
    |      |      |      |      |      |
YTSIPYTVFQ TNSFAERSFC LSL

```

70 Seq ID NO: 92 DNA sequence
Nucleic Acid Accession #: NM_003706.1
Coding sequence: 310-1935 (underlined sequences correspond to start and stop codons)

```

75 1      11      21      31      41      51
    |      |      |      |      |      |
CACGAGGCGG GGGCCATTTT ACCTCCAGGT TGGCCCTGCT CAGGACCAGG AGGAAACACC 60

```

TCCAGCCCCG GACCTCCTCC CACAGGGGGA AAAGGAAGC AGGAGGACCA CAGAAGCTTT 120
 GGCACCGAGG ATCCCGCAG TCTTCACCG CGGAGATTCC GGCTGAAGGA GCTGTCCAGC 180
 GACTACACCG CTAGCGCAG GGAGCCCAAG CCTCCGCACC GGATTCCGGA GCACAAGCTC 240
 CACCGCGCAT GCGCACACGC CCCAGACCCA GGCTCAGGAG GACTGAGAAT TTCTGACCG 300
 5 CAGTGCACCA TGGGAAGCTC TGAAGTTTCC ATAAATCCTG GGCTCCAGAA AGAAGAAAA 360
 GCGGCGGTGG AGAGACGAAG ACTTCATGTG CTGAAAGCTC TGAAGAAGCT AAGGATTGAG 420
 GCTGATGAGG CCCAGTTTGT TGCTGTGCTG GGCTCAGGCG GAGGACTGCG GGCTCACATT 480
 GCTGTCTTGG GGGTCTGAG TGAGATGAAA GAACAGGGCC TGTGGATGC CGTCACGTAC 540
 CTGCGAGGGG TCTCTGGATC CACTTGGGCA ATATCTTCTC TCTACACCAA TGATGGTGAC 600
 10 ATGGAAGCTC TCGAGGCTGA CCTGAAACAT CGATTACCC GACAGGAGTG GGAATGGCT 660
 AAGAGCTTAC AGAAACCAT CCAAGCAGCG AGGTCTGAGA ATTACTCTCT GACCGACTTC 720
 TGGGCTTACA TGGTTATCTC TAAGCAAAACC AGAGAAGCTG CGGAGTCTCA TTGTCCCAAT 780
 ATGAAGAAGC CGGTGAAGA AGGACACTTA CCTTACCCAA TATTTGCAGC CATTGACAAT 840
 GACCTGCAAC CTCTCTGGCA GGAGGCAAGA GCACCAAGAG CTTGTTTGA GTTCAACCCCT 900
 15 CACCAAGCTG GCTTCTCTGC ACTGGGGGCC TTTGTTTCCA TAACCCACTT CGGAAGCAAA 960
 TTCAAGAAAG GAAGACTGGT CAGAACTCAC CCTGAGAGAG ACCTGACTTT CCTGAGAGGT 1020
 TATATGGGAA GTGCTCTTGG TAACACTGAA GTCAATTAGG AATACATTTT TGACCAAGTTA 1080
 AGGAATCTGA CCTGAAAGG TTTATGGAGA AGGGCTGTTG CTAATGCTAA AAGCATTGGA 1140
 20 CCACTTATTT TTGCGGATT ACTGAGGCTG CAAGAAAGTT CACAAGGGGA ACATCCTCCC 1200
 CCAGAAGTGG AAGGCGTGA GCCTGAACAC ACCTGGCTGA CTGAGATGCT CGAAGATTGG 1260
 ACCAGGACCT CCTTGAAGAA GCAGAGGAGG CCCCAGAGG ACCCGAAGG GAAAGGCTCA 1320
 CTCAGTAATC TGATGGAATT TGTGAAGAAA ACAGGCATTT GCGCTTCAAA TTGGGAATGG 1380
 GGGACCACTC ACAACTTCCT GTACAAACAC GGTGGCATCC GGGACAGAT AATGAGCAGC 1440
 CGGAAGACAC TCCACTGGT GGAATGCTGG TTAGCCATCA ACACCTCCTT CCCACTCGTG 1500
 25 CTGCCCCGCA CGCGGAGGT TCACCTCCTC CTCTCCTTGG ACTTCAGTGC CGGAGATCCT 1560
 CTCGAGACCA TCCGGGCTAC CACTGACTAC TCCGCGCGCC ACAAGATCCC CTTTCCCCAA 1620
 GTAGAGAGAG CTGAGCTGGA TTTGTGTGCC AAGGCCCCCG CCAGCTGCTA CATCTGAAA 1680
 GGAGAACTG GACCACTGGT GATACATTTT CCCCTGTTCA ACATAGATGC CTGTGGAGGT 1740
 GATATTGAGG CATGAGTGA CACATACGAC ACATTCAGC TTGCTGACAC CTACACTCTA 1800
 30 GATGTGGTGG TGCTACTCTT GCATTAGCC AAGAAGAATG TCAGGGGAAA CAAGAAGAAG 1860
 ATCTTAGAG AGTTGATGAA CGTGGCCGGG CTCTACTACC CGAAGGATAG TGCCGGAAGT 1920
 TGCTGCTTGG CAAGATGAG CCTCAGCTTC CAGGGCACTG TGGGCTTGT GGTCTACTAG 1980
 GGCCCTGAAG TCCACTGGC CTTCTGTTT TCACTCCCT TCAGCCACAC GCTTCATGGC 2040
 35 CTTGAGTTCA CCTTGGCTGT CCTAACAGGG CCAATCACCA GTGACCAGCT AGACTGTGAT 2100
 TTTGATAGCG TCATTAGAA GAAGGTGTCC AAGGAGCTGA AGGTGTGAA ATTTGTCTG 2160
 CAGGTCCCTC GGGAGATCCT GGAGCTGGAG CATGAGTGTG TGACAATCAG AAGCATCATG 2220
 TCCAATGTCC AGATGGCCAG AATGAATGTG ATAGTTCAGA CCAATGCCTT CCATGCTCC 2280
 TTTATGACTG CACTTCTAGC CAGTAGCTCT GCACAAGTTA GCTCTGTAGA AGTAAGAACT 2340
 40 TGGGCTTAAA TCAATGGGCTA TCTCTCCACA GCCAAGTGA GCTCTGAGAA TACAACAAGT 2400
 GCTCAATAAA TGCTGTCTGA TTGACTGATG AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2460
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA

Seq ID No: 93 Protein sequence

Protein Accession #: NP_003697.1

1 11 21 31 41 51
 50 MGSSEVSIIP GLQKEEKAIV ERRRLHVLKA LKGLRIEADP APVVAVLGSQ GGLRAHIACL 60
 GVLSEMEQGG LLDAVTYLAG VSGSTWATSS LYTNDGDMEA LEADLKHPTP RQEWDLAKSL 120
 QKTIQAAESE NYSLTDFFWAY MVISKQTRRL PESHLSNMKK PVBEGLFVPP IFAAIDNDLQ 180
 PSWQEARAPE TWPEPTPHHA GFSAAGAPVS ITHFGSKFKK GRIVRTHPER DLTFLRLGLWG 240
 55 SALGNTEVIR EYIFDQLNL TLKGLWRRAV ANAKSIGHLI FARLLRLQES SQGEHPFPED 300
 EGGPEHTWL TEMLENWTRT SLEKQEPHE DPERKGSLSN LMDFVKKTGI CASKNEWGTT 360
 HNFLYKIGGI RDKIMSSRKH LHLVDAGLAI NTPLPLVLP TREVHLILSF DFGAGDPFET 420
 IRATTDYCRH HKIPFPQVEE AELDLNSKAP ASCYILKGET GPVVIHFPLF MIDACGGDIE 480
 AWSDTYDTFK LADTYTLDVV VLLALAKKN VRENKKILR ELNMFAGLYY PKDSARSCTL 540
 A

Seq ID No: 94 DNA sequence

Nucleic Acid Accession #: AK027351

Coding sequence: 1-642 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 65 AGGGAAAAAA ACTCCATTAA AAAGCCGAGC TTTCTCCAT GTTAGATGTG ACTTGAAAA 60
 TGAGAAAGAT TTAGCAAAAT TCCACCGTAT CTTTGTCCAG GCTAGAGACA GGGAGAGCAG 120
 70 AGTAAACCCC TCAGGCTGCT GAAATTTCTA GGCTGTTAGG AAGCCCTCG AATTCTGTGA 180
 AAATGAGGGT TTCTTAACCT ACACCTGAGG CGGAAAGGGG CAGACCTCTT TCATAACTCC 240
 CTAAGATGTG TGTATCCTTT CTTTACCAGC ATGGTAAGCA ACAGGACATA TCCCGGCTC 300
 GGACATGTCT GTATGATCCA AGGTACCCAA AGTCAGACAG AGTAACCTCA AGCCTGGCAC 360
 75 TGGCTTCTG CCGCTTCATG TGCTTTGGAA AAAGCAGGAG AAGCAATAGC AGCAGGAGTC 420
 CCCAGCAGCT GGAAGCCGAA GAATGAAGCT CAAAGAGGGA ACTGACAGCA GCTGCGGCTG 480
 CAGGGGCAAC GACGAGAAGA AGATGTTGAA GTGTGTGTGT GTGGGGAGCG GTGCCGTGGG 540

5
 10
 15
 20
 25
 30
 35
 40
 45
 50

GAAAACCTGC CTGCTGATGA GCTACGCCAA CGAGCGCTTC CCAGAGGAAT ACGTGCCAC 600
 TGTGTTGAC CACTATGCAG TTAATGTGAC TGTGGGAGGC AAGCAACACT TGCTGGGACT 660
 GTATGACACC GGGGACAGG AGGACTACAA CCAGCTGAGG CCACCTCTCT ACCCAACAC 720
 GGATGTGTTT TTGATCTGCT TCTCTGTCTT AAACCTGCCC TCTTACCACA ATGTCCAGGA 780
 GGAATGGGTC CCCGAGCTCA AGGACTGCAT GCCTCAGGTG CCTTATGTCC TCATAGGGAC 840
 CCAGATTGAT CTCCGTGATG ACCCAAAAC CTGGGCGCGT TTGCTGTATA TGAAGAGAA 900
 ACCTCTCACT TACGAGCATG GTGTGAAGCT CGCAAAAGCG ATCGGAGCAC AGTGCTACTT 960
 GGAATGTTC GCTCTGACTC AGAAAGGTCT CAAAGCGTCT TTTGATGAAG CAATCCTCAC 1020
 CATTTTCCAC CCCAAGAAA AGAAGAAACG CTGTCTGAG GGTCCAGCTG GCTGTTCAT 1080
 TATCTGAGGT TGTCTGGGAC CTGCTCCAC CCATCCAGG GATGAGAAAT GCAGCCATC 1140
 TCTGTGGCCA AGCTCCAGCC AAAAGGAGG GCAAGACAG AAAGGAACTC CCTTTGCAAG 1200
 GAGGCTTGCC CATCAACCT CTGAGCCCTC CCAACACAG ACATAGTCA GCCACTGCC 1260
 AGACCTCCC TGCCAGCCAG AAGCATCCGT ACTGCAAGCT GTCTGAGAAT GCTGGGCGT 1320
 GATTGACAGC AGTCCCGCTG CTGATGTCAT CAAAACAAA GTCAAGAGCC ATCTCACAT 1380
 TTACAATCC CCAGCTCATG AACGTGAAGC TGATAGGAAA TCACCCAGG GAACCGAAA 1440
 AAGAACTTG ATCTCTCTAT TGCTGGCGTT ACTTGATGTC TTTTATAAAA CTGGGAGCTA 1500
 CAATACTAAC CTTTCTTCTT GAATCTGCTG TTCTACCAT GTGCTCTACA TTCAATTGTA 1560
 TTAATCAAG AATCTACTA ATTTCCAGTT CACTCAGGCC TTACTAATCC ATACCAAAAT 1620
 AGCCTAAGCA CAAGGCATTT TATAITCAT TCTATTTTCA GCATGTTTCT ACCAAGGCTA 1680
 TTAGAACCAA CAGTACCTC TGAATGCCG ATTAAGAA GACATGAGAA GACTTTAAAA 1740
 GTTTTGAAGA CTTTCAAGAG CATGATTTT GAACCTAAT GAAAGAAAAC CATCTGAAT 1800
 GTTCAGGTC CACATTTTTC CCAAGATAC ACTCTATAGA TGCTTAGTAG TGGCTGATT 1860
 TTTTCCATG TATGCCAAG ACAACTAAA AATGAAGTGT GTTGAAGT GTAGTATTT 1920
 TGTATTTTAT CCAAGTTGAT TGGGGGAGA ATATGGCAGG ATCCATCTTT TACAGTATT 1980
 25 TGTATTCAGT AAGTGGACA TTCTGTCTCC TCCCTTCCCC CATTCATGCT CCTCTTCTCT 2040
 CCTTATTTTC ACTTTCTCTC ATGCCCGGAT CTTTTTATTC TCCCAGTTA TAACCCAGTT 2100
 ATAAAGAAA GATCTGAGCA TAAAGATACG TGTTTAAAAA TAACATAAAG TAAAGGAAG 2160
 TGCCCTAATT TTCTATTTG CTTCACCTGA AAGTGCTTCT CAGCTCGCCC CATGTAAGTT 2220
 CTCATTCCAT GTAAATGACA TTTTCCAGTT ACAACTGGTA CTGAGATTTT GCCTCTCTCT 2280
 30 TTTCTACTC ATCTCCCAA ATGTCTTTGT GGGAGCCATA TCAGTGGATA CCAAGCTCTG 2340
 TATCCATTG TCCCTGCGCC TCCACAATGT GTGACATAGA ACAGGAGCTT TGGCGCTGGG 2400
 AAGCTAAGAG CTCCTAGTAA GGAATCTGT GCCCAATGAT GTAAACAAT TCCAAACATC 2460
 CAGGAATTTT TGTATCATAG AGCGAATTAC TTCTATCTTT TTCAATAGAG GCTATGAGGA 2520
 CTTCAATTA TCTTGTAGTG CTTATAAGTG CCTGGAATC ACCAGTAGT GCATTAAT 2580
 35 TTTTCTCAG TTGCTGAGC AAGTCTTTC TTATAGTGT GAAATTAACA CAATTTAAG 2640
 ACTTCCAGA TTCAAGCTCC CACTGTGGA AAAAGCCAGC CTTTCTAATC TCTCTGCTA 2700
 CTGAATAAG CACTTAAGAA TTGGGTGATA GCCAGGCACC GTGGCTCATG CCTGTAATCC 2760
 CAACACTTAG GGAGGCTGAG GTGGGTGGG CGCTTGAGCT CAGGAGTTCA AGACAGCCT 2820
 40 GGGTAATATA GTGAGATCCT GTGTCTCTAT AAAAAAATA AAAATTAGTC AGTTGTAGTG 2880
 ACACATACCT GTAGTCCAG CTAATCAGGA GGCTGAGGTG GAAGGATCAC TTGAGCCAG 2940
 AAGCTAAGGC TGCAGTGAGC TGTGACTGT CCACTACACT CCAGCCTGAG TGACAGAGAA 3000
 AGAATCTGTC AAAAAAATA AAAAAACAAC CTACATTTC AGTACTATT CCCTCTCTCT 3060
 CCACTTAATT GCTAAGATT TTCTTTTATA CGCACACACT CCAGTGACTG GAAAGACGGG 3120
 45 AGTTCTCAGT CAAAGCTTGA CATTTAGAGA AAACAAGGAC TTTCTGCTTT TATAAATGGA 3180
 AATCAACTGT GTATGAAGT TAACTCTGCA GAGGTTATGA ATTCACTCTT TACAACAAT 3240
 AATGAAGTTT TAGTCTGTGA ATAAATGAAA TGTATTAGG CAGCTTTGTT GCATGATTGC 3300
 ATAGTTATAT CTGTCTAAG GGCCTACTAT TTCTCACTGA TGTGATGAA AAAATGAGAG 3360
 CAGTATGTTT CAGGTGTGT GCACCAACA GGCATAAGC TCCGAGGTC ACCACTTCCC 3420
 50 TAATGGGCA CAGGAAGTAA GTTGTCTTG ATGGGAGAT CACGTCACCC AGAACAGCA 3480
 ACTGGATAGA GACTGTGTT AGTGCTCTGG TAGAGCACAG GCTCCAGGG GTCTTAAGAG 3540
 CTAATTAAGT AATAAACAA TCTAGAACA AGCAA

Seq ID No: 95 Protein sequence:
 Protein Accession #: CAC06611.1

55
 60
 180

1 11 21 31 41 51
 MNCKEGTDSS CGCRGNDEKK MLKCVVVGDG AVGKTCLLMS YANDAFPEBY VPTVFDHYAV 60
 TVTVGGKQHL LGLYDTAGQE DYNQLRPLSY PNTDVFLICF SVVNPASYHN VQSEWVPELK 120
 DCMHPVYVVL IGTQIDLRDD PKTLARLLYM KEKFLTYEHG VKLAKAIGAQ CYLECSALTQ 180
 KGLKAVPDEA ILTIHPKKK KKRCSEKHSC CSII

65
 Seq ID NO: 96 DNA sequence
 Nucleic Acid Accession #: NM_003654.1
 Coding sequence: 367-1602 (underlined sequences correspond to start and stop codons)

70
 75

1 11 21 31 41 51
 GGGGAGGGCG CGGAGGGCGG AGGATGCCGC CGCGGCTGCT GCGCGCGCGG CCACCCGCGG 60
 GTCCCGGCGG ACCCTACTCC AGACCGGAGG ATGGAGCGCG CGCTGGGCGC TGCACTGCT 120
 CCCGGGCGCT CCCGACCAAG GTAGCTGGTG TCACTTGGGT GTGGTTGGAA GAAGACTTTC 180
 TCCCGAGCTG CATTCGCGGA GCGCGCCTTT CGACTGGAG GCGGGTCTG CTGGCCACAG 240
 75 GGTGCGCGCA CTGGCTGGGA CTGCGAGCTG GGCCTGGAGA CGCTGGTGGC TGTGGACTCC 300
 CCAGCTTGGG GCAGTCCCTC TTTGACCTCA CCCCTGGAG AAGCAGCCCC ATGAAGGTGC 360

5
 10
 15
 20
 25
 30
 35

CCAGCCATGC AATGTTCTCTG GAAGGCCCTC CTCTCTCTTG CCCTGGCCTC CATTGCCATC 420
 CAGTACACGG CCATCCGACAC CTTCACGCGC AAGTCTCTTC ACACCTGCCC CGGGCTGGCA 480
 GAGGCGCGGC TGGCCGAGCG ACTGTGCGAG GAGAGCCCA CCTTCGCTTA CAACCTCTCC 540
 CGCAAGACCC ACATCCTCAT CCTGGCCACC ACGCGCAGCG GCTCTCTCTT CGTGGGCCAG 600
 CTCTTCAACC AGCACCTGGA CGTCTTCTAC CTGTTTGAAG CCCTCTACCA CGTCAGAAC 660
 ACGCTCATCC CCGCTTCTAC CCAGGGCAAG AGCCCGGCGG ACCGGCGGGT CATGCTAGGC 720
 GCCAGCGCGC ACCTCTCTCG GAGCCTCTAC GACTGCGACC TCTACTTCTT GGAGAACTAC 780
 ATCAAGCGCG CGCGGCTCAA CCACACCAAC GACAGGATCT TCCGCGCGGG GGCCAGCGGG 840
 GTCTCTGCTT CCGGCGCTGT GTGCGACCTT CCGGCGCGAG CGGACCTGGT CTTGGAGGAG 900
 GGGGACTGTG TGCGCAAGTG CGGCTACTTC AACCTGACCG TGGCGGCCGA GGCCTGCCGC 960
 GAGCGCAGCC ACCTGGCCAT CAAGACGGTG CGCTGCGCG AGGTGAACGA CCTGGCGGCC 1020
 CTGCTGGAAG ACCGCGGATT AAACCTCAAG GTCATCCAGC TGGTCCGAGA CCCCAGCGGC 1080
 ATTCTGGCTT CGCGCAGCGA GACCTTCCGC GACACGTACC GGCTCTGGCG GCTCTGTATC 1140
 GGCACCGGGA GGAAACCTTA CAACCTGGAC GTGACGACGC TGACACCGGT GTGCGAGGAC 1200
 TTCTCAACTC CCGTGTCCAC CGGCTCATG CGGCCCTCGT GGCTCAAGGG CAAGTACATG 1260
 TTGCTGCGCT ACGAGGACCT GGCTCGGAAC CCTATGAAGA AGACCGAGGA GATCTACGGG 1320
 TTCTCTGGAG TCCGCTGGA CAGCCAAGTG GCCCGCTGGA TCCAGAACAA CACGCGGGGC 1380
 GACCCACCCG TGGCAAGCA CAATACGGC ACCGTGCGAA ACTCGCGGCG CACGCGCGAG 1440
 AAGTGGCTCT TCCGCTCTC CTACGACATC GTGGCTTTG CCGCAAGCGC CTGCGCAGAG 1500
 GTCTGCGCTC AGCTGGGCTA CAAGATCGCC GCTTCGAGAG AGGAGCTGAA GAACCCCTCG 1560
 GTCAGCTGGG TGGAGGAGCG GAGCTTCCGC CCCTTCTCGT GACCCGCGCG GTGCGGGTGG 1620
 GGGCGGGAGG CGCAGGCTGT CGGTTTGTAT AAAATGGACC GTTTTAACT GTTGCCTTAT 1680
 TAACCCCTCC CTCTCCACCC TCATCTTGT GTCTTCTCT CCCCAGCTC ACCCCACTCC 1740
 CTCTGACCCC TTTTGTCTCT CTGAAATTG CACTACGTCT TGGACGGGAA TCACTGGGGC 1800
 AGAGGGCGCC TGAAGTAGGG TCCGCGCCCC CCCACCCAT TCACACACAT GGATGTTGGG 1860
 TCTCTGTGGG GACGGTGACA ATGTTTACAA GCACACATT TACACATCCA CACACGACA 1920
 CGGCACTCGC CGAGGCGACT TCTCAAGCTT TTGAATGGGT GAGTGGTGG GTATCTAGTT 1980
 TTGCACTGT CTACTATTG AAGGTAAGAG GATACAAACA AGAGGACAC TTTGCTCTAA 2040
 TTTATGAATG GTGTCCATCC TTTCCCATC CCGCTCTCT GCGCTGACG CCACTTTCCC 2100
 CCCTTAGAGC AGCGAAACTG CCCCCTCTG CCGGCTCTG CTGTGCGTG AGGCAGGTTT 2160
 TTACTGTGAG GTGAACGTGG ACCTGTTTCT GTTTCCAGTC TGTGGTGATG CTGTCTGTCT 2220
 GTCTGAGTCT CGTGGCGGCC CCGTGACCAG TGATGACTGA TGAATCTTAT GAGCTCTCTA 2280
 TTGATCTCGG GGTCCATCTG TGATATTCT TTGTGCCAAA AAGAAAAAAA AAGAGTGGAT 2340
 CAGTTTGCTA AATGAACATT GAAATTGAAA TGCTTTATCT GTGTTTCTG TAAATAAAG 2400
 AGTCAATAA TCACC

Seq ID No: 97 Protein sequence:
 Protein Accession #: NP_003645.1

40
 45
 50

1 11 21 31 41 51
 MQCSWKAVLL LALASIAIQY TAIRTFYAKS FHTCPGLASA GLABRLCEES PTFAYNLARK 60
 THILLIATTR SSSSPVQQLF NQHLDFVYLF EPLYHVQNTL IPRFTQKESP ADRRVMLGAS 120
 RDLLRLDYDC DLYFLENYIK PPPVNHITDR IPRRGASRLV CSRFPVCDPPG PADLVLEBGD 180
 CKRCKGLNL TVABACRER SHVAIKTVRV PEVNDLRALV EDPRILNLKI QLVDRPRLIL 240
 ABRSEFTFDD YRLWRLWYGT GRKPYMLDVT QUITVCEDFS NSVSTGLMRP PWLKGKMYLV 300
 RYEDLARNPM KTEBIYGFLL GIPLDSHVAR WIQNNTGDDP TLGKHKYGTV RNSAATAEKM 360
 RFLRLSDIVA FAQNACQQLV AQDGYKIAAS EEELKNPVS LVEERDFRFP S

Seq ID NO: 98 DNA sequence
 Nucleic Acid Accession #: NM_002852.1
 Coding sequence: 68-1213 (underlined sequences correspond to start and stop codons)

55
 60
 65
 70
 75

1 11 21 31 41 51
 CTCAAACTCA GCTCACTTGA GAGTCTCTCT CCGCCAGCTG TGGAAGAAGC TTTCGCTCTC 60
 TCCAGCAATG CATCTCTCTG CGATCTCTGT TTGTGCTCTC TGGTCTGAG TGTGGCCGA 120
 GAACTCGGAT GATTATGATC TCATGTATGT GAATTTGGAC AACGAAATAG ACAATGGACT 180
 CCATCCCACT GAGGACCCCA CGCGGTGCGA CTGCGGTGAG GAGCACTCGG AATGGGACAA 240
 GCTCTTCATC ATGCTGAGAG ACTCGCAGAT GAGAGAGCGC ATGCTGCTGC AAGCCACGGA 300
 CGACGTCCTG CCGGGCGAGC TGCAGAGGCT CCGGGAGGAG CTGGGCGCGC TCGCGGAAG 360
 CTTGGCGAGG CGGTGCGCGC CGGGGGCTCC CGCAGAGGCC AGGCTGACCA GTGCTCTGGA 420
 CGAGCTGCTG CAGGCGACCC GCGACGCGGG CCGCAGGCTG CGCGGTATGG AGGGCGCGGA 480
 GCGCGAGCGC CCAGAGGAGG CGGGGCGCGC CTTGGCGCGG GTGCTAGAGG AGCTGCGGCA 540
 GACGCGAGCC GACCTGCACG CGGTGCAAGG CTGGGCTGCC CGGAGCTGGC TGC CGGCAGG 600
 TTGTGAACA GCTATTTTAT TCCCAATGCG TTCCAAGAAG ATTTTGGGAA GCGTGCATCC 660
 AGTGAGACCA ATGAGGCTTG AGTCTTTTAG TGCTGCAATT TGGGTCAAG CCACAGATGT 720
 ATTAACAAA ACCATCCTGT TTTCTTATGG CACAAAGAGG AATCCATATG AAATCCAGCT 780
 GTATCTCAGC TACCAATCCA TAGTGTTTGT GGTGGGTGGA GAGGAGAAC AACTGGTTGC 840
 TGAAGCCATG GTTTCCTTGG GAAGGTGAC ACACCTGTGC GGCACCTGGA ATTCAAGGGA 900
 AGGCTCACA TCCTTGTGGG TAAATGTGTA ACTGGCGGCT ACCACTGTTG AGATGGCCAC 960
 AGGTCACTT GTTCTCTGAG GAGGAATCCT CGAGATTGGC CAAGAAGAAG ATGGCTGCTG 1020
 TTTGGTGGT GGCCTTGTAT AACATTAGC CTCTCTGGG AGACTCACAG GCTCTCAATAT 1080
 CTGGGATAGT GTTCTTAGCA ATGAAGAGAT AAGAGAGACC GAGAGGACAG AGTCTGTCTA 1140

CATCCGGGGG AATATTGTTG GGTGGGGAGT CACAGAGATC CAGCCACATG GAGGAGCTCA 1200
 GTATGTTTCA TAAATGTTGT GAAACTCCAC TTGAAGCCAA AGAAGAAAC TCACACTTAA 1260
 AACACATGCC AGTTGGGAAG GTCTGAAAAC TCAGTGCCATA ATAGGAACAC TTGAGACTAA 1320
 TGAAGAGAG AGTTGAGACC AATCTTTATT TGTACTGGCC AATACTGAA TAAACAGTTG 1380
 AAGGAAAGAC ATTGGAAAAA GCTTTTGAGG ATAAATGTTAC TAGACTTTAT GCCATGGTGC 1440
 TTTTCAGTTA ATGCTGTGTC TCTGTGAGAT AAACCTCTCA ATAAATTAATA AGGACTGTAT 1500
 TGTGTAACAG AGGGAACAAT GTTTTACTTT TCTTTGGTTA ATTTTGTTTT GGCCAGAGAT 1560
 GAATTTTACA TTGGAAGAAT AACAAATATA GATTTGTGT CCATGTTTCA TTGTTATTGG 1620
 TATGTACCTT ATTACAAAAA AAATGATGAA AACATATTTA TACTACAAGG TGACTTAACA 1680
 ACTATAAATG TAGTTTATGT GTTATAATCG AATGTCAGT TTTTGAGAG ATAGTCATAT 1740
 AAGTTATATT GCAAAAGGGA TTTGTATTAA TTTAAGACTA TTTTGTGAAA GCTCTACTGT 1800
 AATAAATA TTTTATAAAA CTAATAAAAA AAAAAA

Seq ID No: 99 Protein sequence:
 Protein Accession #: NP_002843.1

1 11 21 31 41 51
 20 MHLAILFCA LMSAVLAENS DDYDLMYVNL DNEIDNGLHP TEDPTPCDCG QHSEWDKLF 60
 IMLEKSMRE RMLLQATDDV LRGLQLRLR ELGLRLAELA RCPAPAPAE ARLTSALDEL 120
 LQATRDAGRR LARNGAENQ RPEAGRALA AVLEBLQTR ADLHAVQWA ARSWLPAGE 180
 TAILPFMRK KIPGSVHPVR PMRLSFSAC IWKATDVLM KTLFSYGTG RNPYSIQLYL 240
 SYQSIIVPVG GERNKLVARA MVLGRWTHL CGTWSEBGL TSLWVNGELA ATTVMATGH 300
 25 IVPBGGLLQI QKRGKCCVG GGFDETLAPS GRLTGPNIND SVLSNREIRE TGGAESCHIR 360
 GNIVGWVTE IQPHGGAQYV S

Seq ID NO: 100 DNA sequence
 Nucleic Acid Accession #: NM_007351.1
 30 Coding sequence: 72-3758 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 35 CTGCTATCAA AAAGGCCATA AGGATTTTGT CCCCAATTT CACATGAGCT ACCTTGCTTC 60
 AAACACTGTA GTGAAGGGG GCAAGATTAT TTGTCCTTCT TTCTAGTTTA TGGAGTGGGG 120
 GCATTTGGCT TAACAACAGT AAGCATTTCT GGACTATACC TGAGGATGGG AACTCTCAGA 180
 AGACTATGCC TTCTGCTTCA GTTCTCCCAA ATAAATACA AAGTTTGCAA ATACTGCCAA 240
 40 CCATCGGGT CATGTCGGG GAGATAGCTA CAATCCAGA GGCAAGAACT TCTGAAGACA 300
 CTCTCTTTAA ATCAACACTG CCTCCCTCAG AAACAAGTGC ACCTGCTGAG GGTGTGAGAA 360
 ATCAAACTCT CACATCCACA GAGAAAGCAG AAGGAGTGGT CAAGTACAG AATCTTACCC 420
 TCCCAACCAA CGCTAGCATC AAGTTCAATC CTGGAGCAGA ATCAGTGGTC CTTTCCAATT 480
 CTACACTGAA ATTTCTTCAG AGCTTTGCCA GAAAGTCAAA TGAACAAGCA ACTTCTCTAA 540
 45 ACACAGTTGG AGGCACTGGA GGCATTGGAG GCGTTGGAGG CACTGGAGGC GTGGGAAATC 600
 AGCCCCACG GGAACATAC CTCAGCGGG GTGACAGCAG TTCCAGCCAA AGAACTGACT 660
 ACCAAATATC AAATTTGCAA ACAAAGTAGG GAAAGAATTG GTGTGCTTAT GTACATACCA 720
 GGTATCTCC CACAGTGACA TTGGACAACC AGGTCACTTA TGTCCAGGT GGGAAAGGAC 780
 CTGTGTGCTG GACGGTGGGA TCCTGTCTCT AGAGATCTCA GAAGATATCC AATCCTGTCT 840
 50 ATAGGATGCA ACATAAAATT GTACCTCAT TGGATTGGAG GTGCTGTCTT GGATACAGTG 900
 GGCCGAATG TCAACTAAGA GCCCAGGAAC AGCAAGATTG GATACACACC AACCAGGCTG 960
 AAAGTCATAC AGCTGTTGGC AGAGGAGTAG CTGAGCAGCA GCAGCAGCAA GGCTGTGGTG 1020
 ACCCAGAAGT GATGCAAAAA ATGACTGATC AGGTGAACTA CCAGGCAATG AAAGTACTC 1080
 55 TTCTCAGAAA GAAGATTGAC AATATTCTTT TGAATGTGAA TGAATGAAG AACACTTACT 1140
 CCTCCCTAGA AGGAAAAGTC AGCGAAGATA AAGCAGAGA ATTTCAATCT CTCTTAAAG 1200
 GTCTAAATAT CAAAAGCATT AATGTACTGA TAAGAGACAT AGTAAGAGAA CAATTTAAAA 1260
 TTTTTCAAAA TGACATGCAA GAGACTGTAG CACAGCTCTT CAAGACTGTA TCAAGTCTAT 1320
 CAGAGACCT CGAAGCACC AGGCAATATA TTCAAAAAGT TAATGAATCT GTGGTTTCAA 1380
 TAGCAGCCCA GCAAAAGTTT GTTTTGGTGC AAGAGAATCG GCCACTTTG ACTGATATAG 1440
 60 TGAACATAAG GAATCACATT GTGAATGTAA GGCAAGAAAT GACTCTTACA TGTGAGAAGC 1500
 CTATTAAAGA ACTAGAAGTA AAGCAGACTC ATTTAGAAGG TGCTTAGAAA CAGGAACACT 1560
 CAGAGAAGCT TCTGTATTAT GAATCCCTCA ATAAAACTCT TTCTAAATTG AAGGAAGTAC 1620
 ATGAGCAGCT TTTATCAACT GAACAGGTAT CAGACCAGAA GAATGCTCCA GCTGCTGAGT 1680
 CAGTTAGCAA TAATGTCACT GAGTACATGT CTACTTTACA TGAATAATA AAGAAGCAGA 1740
 65 GTTTGATGAT GCTGCAATG TTTGAAGATT TGCAATTTCA AGAAGCAGG ATTAACAATC 1800
 TCACCGTCTC TTTGAGATG GAGAAAGAGT CTCTCAGAGG TGAATGTGAA GACATGTTAT 1860
 CCAATGCGAG AAATGATTTT AAATTTCAAC TTAAGGACAC AGAAGAGAAT TTACATGTGT 1920
 TAAATCAAAC ATTGGCTGAA GTTCTCTTTC CAATGGACAA TAAGATGGAC AAAATGAGTG 1980
 AGCAACTAAA TGATTTGACT TATGATATGG AGATCCTTCA ACCCTTGCTT GAGCAGGGAG 2040
 70 CATCACTCAG ACAGACAATG ACATATGAAC AACCAGAGGA AGCAATAGTG ATAGGGAATA 2100
 AGATAGAAAA TCTGACTAGT GCTGTCAATA GTCTAAATTT TATTATCAAA GAACCTTACAA 2160
 AAGACACAAA CTTACTTAGA AATGAAGTAC AGGTCGTGTA TGATGCCTTA GAAGAAGCTA 2220
 TCAATTAATA TGCTTAGAAA ATGGAAGATG GCCTCAATAA GACAATGACT ATTATATAA 2280
 75 TTGCTATTGA TTGCTATTCA GATAACTATG CCTTAAAGA GACTTTAAGT ACTATTAAAG 2340
 ATAATAGTGA GATCATCAT AAATGTACCT CCGATATGGA AACTATTTTG ACATTTATTC 2400
 CTCAGTTCCA CCGTCTGAAT GATTCATATC AGACTTTGGT CAATGACAA CAGAGATATA 2460

5 ACTTTGTTTT GCAAGTCGCC AAGACCCCTTG CAGGTATTC CAGAGATGAG AAATAAATC 2520
 AGTCCAACTT CCAAAAGATG TATCAAAATOT TCAATGAAAC CACTTCCCAA GTGAGAAAT 2580
 ACCAGCAAAA TATGAGTCAT TTGGAAGAAA AACTACTCTT AACTACCAAG ATTTCCAAA 2640
 ATTTTGAGAC TCGGTTGCAA GACATTGAGT CTAAGTGTAC CCAGACGCTC ATACCTTATT 2700
 ATATTTCAGT TAAAAAGGC AGTGTAGTTA CAAATGAGAG AGATCAGGCT CTCTACTGTC 2760
 AAGTATTAAA TTCCAGATTT AAGGCGTTGG AAGCAAAATC TATCCATCTT TCAATTAACT 2820
 TCTTTTCGCT TAACAAAACCT CTCCAGCAAG TTTTAACAAT GTGTCACAAT GCTTCTACAA 2880
 GTGTGTCAGA ACTGAATGCT ACCATCCCTA AGTGGATAAA ACATTCCCTG CCAGATATTC 2940
 AACTTCTCTA GAAAGGTCTA ACAGAAATTT TGGAAACCAAT AATTCAAATA AAAACTCAAG 3000
 10 CTGCCCTATC TAATTCAACT TGTGTATAG ATCGATCGTT GCGTGTAGT CTGGCAAATG 3060
 TTGTCAAGTC TCAGAAAGCAA GTAAATCMT TCCCAAGAA AATTACCGCA CTTAAGAAAC 3120
 CAACGGTAAA TCTTACCACA GTCTGTATAG GCGGACTCA AAGAAACAGG GACAACATA 3180
 TATATCCCTA GAGTATTCA AGCTGTAGTC GGCATCCGTG CCAAAATGGG GGCAGTGCA 3240
 TAATGTGAAG AACTAGCTTT ACCTGTGCTT GCAGACATCC TTTTACTGGT GACAACGTCA 3300
 15 CTATTCAAGCT TGTGGAAGAA AATGCTTTAG CTCCAGATTT TTCCAAAGGA TCTTACAGAT 3360
 ATGCACCCAT GGTGGCATTT TTTGCATCTC ATACGTATGG AATGACTATA CCTGTCCCTA 3420
 TCGTGTAAA TAACITGGAT GTCAATTATG GAGCTTCATA TACCCCAAGA ACTGGAATA 3480
 TAGAATTCCT GTATCTTGGG GTATATGTTT TCAAGTACAC CATCGAGTCA TTTAGTGCTC 3540
 ATATTCTTTC ACTTTTATGT GTTATGGA TAGACAAGCT TGCATTTGAG TCTGAAATA 3600
 20 TTAACAGTGA AATACACTGT GATAGGGTTT TAACITGGGA TGCTTATTA GAATTAAAT 3660
 ATGGGCAGGA AGTCTGGTTA CGACTTGCAA AAGGAACAA TCCAGCCAG TTTCCCTG 3720
 TTACTACATT TAGTGGCTAT TTATTATATC GTACATAAGT TAGTATGAAA AACAGACTAT 3780
 CACCTTTATT GAGAAACAGC CAGTGTTTTC ATTTATCTTT GCTTGCACAT CTGCTCTGTT 3840
 25 TTGTTTTTTC TACAGGAAAT GAAATCAAC TTGTTTTTTC AATATGAGTA AACTGTGATG 3900
 TCTATTATAT AAAATTATTT GAATATTGTT TAATGTCTGA ATATGAAAGA GTTCTTGATC 3960
 CTAAAGAAAT TTAGTGGCAC AGAAACAAA GTGAATTGTT TAGCATAATT ATTCTATTC 4020
 TTATTCTTTC ATTTTAAGTC ATTGCAATGG AAGTAATAT TATAAACCGG TAATTACAAC 4080
 ATATTATCAG TCACAGTTTT CTTCCTAATT AAACACTTAA CTTTGTATAT TCCCTGTATA 4140
 30 TAAATATATA ACACACATTT TCTAGATTCA CAAATTAAA TAAATTACTC AAAAAATG

Seq ID No: 101 Protein sequence
 Protein Accession #: NP_031377.1

35 1 11 21 31 41 51
 MKGARLFVLL SLSWSGGIGL NNSKHSWTIP HDGNSQKTMF SASVPPNKIQ SLQLIPTTRV 60
 MSABRIATTFE ARTSEDLLK STLPPSETSA PABGVNRQTL TSTEKAEVGV KLQNLTLPTN 120
 ASIKFNPQAE SVVLSNSTLK FLQSPARKSN EQATSLNVTG GTGGIGGVGG TGGVGNRAPR 180
 40 ETYLSRGDS SSGRTDYQKS NFETTRGKNW CAYVHRLSP TVTLDNQVTV VPGGKPGCW 240
 TGGSCPQRSQ KISNPVYRMQ HKIVTSLDNR CCPGYSGPKC QLRAGEQQLS IHTNQAESHT 300
 AVGRGVABQQ QQQGGDPFV MQKMTDQVNY QAMKLTLLQK KIDNISLTVN DVRTYSSLE 360
 GKVSSEKSRV FQSLKGLKLS KISNVLIRDI VREQFKIPON DMQETVAQLF KTVSSLSDEL 420
 ESTRIIIOKV NESVSVIAAQ QKPVLVQENR PTLTDIVELR NHIVNVRQEM TLTCEKPIKE 480
 LEVQKTHLEQ ALRQEHRSRI LYYESLNKTL SKLKEVHEQL LSTQVSDQK NAPAESVSN 540
 45 NVTEYNSTHL ENIKKQSLMM LQMFEDLHIQ ESKINHLTVS LEMEKESLRG ECEMDLSKCR 600
 NDTPKFLKDT EENLHVLNQI LAEVLFPMDN RMDKMSQLN DLTVDMEILQ PLLEQGASLR 660
 QTMVYBPQKE AIVIRKKIEN LTSAVNSLNF IIKELTKRHN LLRNEVQGRD DALERRINBY 720
 ALEMEDGLNK TWTIINNAID FIQDNVALKE TLSTIKDNSE IHKCTSDME TILTFIPQPH 780
 50 RLNDSTQTLV NDNQRYNFVL QVAKTLGAP RDEKLNQSNF QMVMQMFNET TSQVRYQQQN 840
 MSHLESKLLL TTKISKNFET RLQDIRESKV QTLIPYYSV KRGSVVTNER DQALQLQVLN 900
 SRFLKALEAKS IHLISNPFSL NKTLEHVLTM CHNASTVSSE LNATIPKWKI HSLPDIQLLQ 960
 KGLTEFVEPI IQIKTQAALS NSTCCIDRSL PGLANVVK S QKQVKSPLPK INALKKPTVN 1020
 VESNALAPDF SKGSRYVAPM VAFASHTYG MTIPGRILEN NLDVNYGASY TPRTKGFRIP 1080
 55 YLGVVYFKYT IESPSAHISQ FLVVDQIDKL AFSEENINSE IHCDRLVITGD ALLELWYQGE 1200
 VMLRLAKGII PAKFPVPTTF SGYLLYRT

Seq ID NO: 102 DNA sequence
 Nucleic Acid Accession #: NM_000873.2
 Coding sequence: 57-884 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51
 ATCTCCCTCC AGGCAGCCCT TGGCTGTGCC CTGCGAGCCC GTGGAGACTG CCAGAGATGT 60
 CCTCTTTCGG TTACAGGACC CTGACTGTGG CCTCTTTCAC CCGATCTGTC TGTCCAGGAT 120
 CGGATGAGAA GGTATTGAGG GTACACGTGA GGCCAAAGAA GCTGGCGGTT GAGCCCAAG 180
 GGTCCCTCGA GGTCAACTGC AGCACCACCT GTAACCAAGC TGAAGTGGGT GGTCTGGAGA 240
 CCTCTCTAAA TAAGATTCTG CTGGACGAAC AGGCTCAGTG GAAACATTAC TTGTTCTCAA 300
 70 ACATCTCCCA TGACACGGTC CTCCAATGCC ACTTCACCTG CTCCGGGAAG CAGGAGTCAA 360
 TGAATTCCAA CGTCACGGTG TACCAGCCTC CAAAGCAGGT CATCTGACA CTGCAACCCA 420
 CTTTGTGTGC TGTGGGCAAG TCCTTCACCA TTGAGTGAG GGTGCCACC GTGAGGCCCC 480
 TGGACAGCCT CACCTCTTTC CTGTTCCGTG GCAATGAGAC TCTGCATAT GAGACCTTGG 540
 75 GGAAGGCAGC CCTGTCTCCG CAGGAGGCCA CAGCCACATT CAACAGCAGC GTGACAGAG 600
 AGGATGGCCA CGCAACTTC TCCTGCCTGG CTGTGCTGGA CTTGATGTCT CGCGGTGGCA 660
 ACATCTTCA CAACACTCA GCCCCGAAGA TGTGTGAGAT CTATGAGCCT GTGTGCGACA 720

5 GGCAGATGGT CATCATAGTC ACGGTGGTGT CGGTGTGTCT GTCCCTGTTC GTGACATCTG 780
 TCCTGCTCTG CTTCATCTTC GGCCAGCACT TGCCGCCAGCA GCGGATGGGC ACCTACGGGG 840
 TGGAGCGGC TTGGAGGAGG CTGCCCCAGG CCTTCCGGCC ATAGCAACCA TGAGTGGCAT 900
 GGCCACCACC ACGGTGGTCA CTGGAATCA GTGTGACTCC TCAGGGTTGA GGTCCAGCCC 960
 TGGCTGAAGG ACTGTGACAG GCAGCAGAGA CTGGGACAT TGCTTTTCT AGCCCGAATA 1020
 CAAACACCTG GACTT

Seq ID No: 103 Protein sequence:
 Protein Accession #: NP_000864.1

10
 15
 1 MSSFGYRTLT VALFTLICCP GSDEKVFVHV VRPKKLAVEP KGSLEVNCSY TCNQPEVGGGL 60
 ETSIAKILID EQAQWIKHYLV SNISHDITVLQ CHFTCSGKQE SMNSNVSVYQ PPRQVILTLQ 120
 PTLVAVKQFP TIECRVPTVB PLDSLTLPLF RGNETHHYET PKAAPAPQE ATATPNSTAD 180
 REDGHRNFGC LAVLDLMSRG GNIFPKHSAP KMLEIYFVPS DSQVMIIIVT VSVLLSLFVT 240
 SVLLCFIFGQ HLRQQRMGTY GVRAAWRRLP QAFRP

20 Seq ID NO: 104 DNA sequence
 Nucleic Acid Accession #: NM_001795.2
 Coding sequence: 121-2475 (underlined sequences correspond to start and stop codons)

25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 1 GACGGTCGGC TGACAGGCTC CACAGAGCTC CACTCAGCTC CAGGCCCTGG ACGGACAGGC 60
 AGTCCAAACGG AACAGAAACA TCCCTCAGCC CCACAGGCAC GATCTGTTC TCCTGGGAAG 120
 ATGCAGAGGC TCATGATGCT CCTCGCCACA TCGGGCGCCT GCCTGGGCTT GCTGGCAGTG 180
 GCAGCAGTGG CAGCAGCAGG TGCTAACCTT GCCCAACGGG ACACCCACAG CCGTGTGCC 240
 ACCCAACGGG GCCAAAGAGG AGATTGGATT TGGAAACGGA TGACATTTGA TGAAGAGAAA 300
 AACACCTCAC TTCCCATCA TGTAGGCAAG ATCAAGTCAA GCGTGAATCG CAAGAATGCC 360
 AAGTACCTGC TCAAGAGAGA ATATGTGGGC AAGGTCTTCC GGGTCAATGC AGAGACAGGA 420
 GACGTGTTCG CCAATTGAGG GCTGGACCGG GAGAATATCT CAGAGTACCA CCTCACTGCT 480
 GTCATTTGTG ACAAGGACAC TGGTGAAGAC CTGAGAGCTC CTTCAGCTT CACCATCAAA 540
 GTTCATGAGC TGAAGCAGCA CTGGCCCTGT TTCAAGCATC GGTGTGTCAA TGGTCCGCTG 600
 CTTGAGTCTG CGCTGTGGG GACCTCAGTC ATCTCTGTGA CAGCAGTGGG TGCAGACGAC 660
 CCCACTGTGG GAGACCAAGC CTCTGTCTAT TACCAATATC TGAAGGGGAA AGAGTATTTT 720
 GCCATGATA ATTCTGGAGG TATTATCACA ATAACGAAAA GCTTGGACCG AGAGAAGCAG 780
 GCCAGGTATG AGATCGTGGT GGAAGCGCGA GATGCCAGG GCTCCGGGGG GGAAGTCCGG 840
 ACGGCCACCG TGCTGTTCAC TCTGCAAGAC ATCAATGACA ACTTCCCTCT CTTACCCAG 900
 ACGGATGACA CATTTGTCTG GCTTGAAGAC ACCGCTGTGG GCACCTCTGT GGGCTCTCTG 960
 TTTGTGTAGG ACCCAGATGA GCCCAGAAC CGGATGAACA AGTACAGCAT CTTGGGGGG 1020
 GACTACAGCG ACGCTTTTCA CATTTGAGCA AACCCCGCCC ACAACGAGGG CATCATCAAG 1080
 CCCATGAGCG CTCTGGATTA TGAATACATC CAGCAATACA GCTTCATGCT CGAGGCCACA 1140
 GACCCACACA TCGACCTCCG ATACATGAGC CTTCCCGCGG GAAACAGAGC CAGAGTCATT 1200
 ATCAACATCA CAGATGTGGA CGAGCCCCCC ATTTTCCAGC AGCCTTTCTA CCACTTCCAG 1260
 CTGAAGGAAA ACCAGAAGAA GCCTCTGATT GGCACAGTGC TGGCCATGGA CCGTGATGCG 1320
 GCTAGGCATA GCATTGGATA TCCCATCCGC AGGACCAATG ACAAGGGCCA GTTCTTCCGA 1380
 GTCACAAAAA AGGGGGACAT TTACAATGAG AAAGAAGTGG ACAGAGAAGT CTACCCCTGG 1440
 TATAACCTGA CTGTGGAGGC CAAAGAACTG GATTCCACTG GAACCCCAAC AGGAAAAGAA 1500
 TCCATTGTGC AAGTCCACAT TGAAGTTTGG GATGAGAAAG ACAATGCCCC GGAGTTTGCC 1560
 AAGCCCTACC AGCCCAAGT GTGTGAGAAG GCTGTCCATG GCCAGCTGGT CCGTGCAGATC 1620
 TCOCGAATAG ACAAGGACAT AACACCAAGA AACGTGAAGT TCAAAATTCAC CTTGAATACT 1680
 GAGAACACTT TTAACCTCAC GATATATCAC GATAACAGG CCAACATCAC AGTCAAGTAT 1740
 GGGCAGTTTG ACCGGGAGCA TACCAGAGTC CACTTCTTAC CGTGTGTCAT CTCAGACAAAT 1800
 GGGATGCCAA GTGCAAGGG CACCAAGCAG CTGACCGTGG CGTGTGCAAA GTGCAAGCAG 1860
 CAGGGCGAGT TCACCTTCTG CGAGGATATG GCGGCCAGG TGGGGGTGAG CATCCAGGCA 1920
 GTGGTAGCCA TCTTACTCTG CATCTCACC ATCACAGTGA TCACCTGCT CATCTTCTG 1980
 CGGCGGCGGC TCCGGAAGCA GCGCCGCGGG CACGGCAAGA GGTGCGCGGA GATCCAGCAG 2040
 CAGCTGTGTA CCTACGACGA GGAGGGCGGC GCGGAGATGG ACACCAACAG CTACGATGTG 2100
 TCGGTGCTCA ACTCGGTGCG CCGCGGCGGG GCGAAGCCCC CGCGGCCCGG GCTGGAGGCC 2160
 CGGCTTTCCT TCTATGCGCA GGTGCAAGAG CCACGAGGC ACGGCGCTGG GGCACACGGA 2220
 GGGCCCGGGG AGATGGCAGC CATGATCGAG GTGAGAGAGG ACGAGGCGGA CCAAGCAGGC 2280
 GACGGCCCCC CTACGACAC GCTGCACATC TACGGCTACG AGGGCTCGGA TCCATAGCC 2340
 GAGTCCCTCA GCTCCCTGGG CACCGACTCA TCCGACTCTG ACGTGGATTA CCACTTCTCT 2400
 AACGACTGGG GACCCAGGTT TAAGATGCTG GCTGAGCTGT ACGGCTCGGA CCCCCGGGAG 2460
 GAGCTGCTGT ATTAGGGGCG CGAGGTCACT CTGGGCTTGG GGACCCAAAC CCCTTGAGC 2520
 CCAGGCCAGT CAGACGCCAG GCACCAAGGC CTCCAAAAAT GGCAGTGACT CCCCAGCCCA 2580
 GCACCCCTTC CTCGTGGGTC CCAGAGACCT CATCAGCTTT GGGATAGCAA ACTCCAGGTT 2640
 CTGGAATAT CCAGGAATAT ATGTCACTGA TGACTATTCT CAAATGCTGG CAAATCCAGG 2700
 CCGTGTGCTG GTCTGGGCTC AGACATCCAC ATAACCTGT CACCCACAGA CCGCGTCTA 2760
 ACTCAAGAC TTCTCTGGC TCCCAAGGC TGCAAGCAA AACAGACTGT GTTTAACTGC 2820
 TATCTGCTG GAGGCAAGG TCCCTGAACG CCGTGTGAGG GCTGGTGGG TCCTGGTGCC 2880
 TATCTGCTG GAGGCAAGG CCGTGAAGC TTGACTTGTG GGGCAGGATT CTCTGAGCC 2940
 CATTCCCAAG GGAGACTGAC CATCATGCCC TCTCTGGGA GCCCTAGCCC TGCTCCAAC 3000

5 CCACTACTCCA CTCCAGTGC CCCACCACTC COCAACCCCT CTCCAGGCCT GTCAAGAGGG 3060
 AGGAAGGGGG CCCTAGGGCAG CTCTGACCT TGGGTCTGA AGTGACCTCA CTGGCTGCC 3120
 ATGCCAGTAA CTGTGCTGTA CTGAGCACTG AACCACATTC AGGGAATGG CTTATTAAAC 3180
 TTTGAAGCAA CTGTGAATTC ATTCTGGAGG GGCAGTGGAG ATCAGGAGTG ACAGATCACA 3240
 GGGTGAAGGC CACCTCCACA CCCACCCCTT CTGGAGAAGG CCTGGAAGAG CTGAGACCTT 3300
 GCTTTAGAGC TCCTCAGCAC CCTCCAGTT TTGCTGAGA AGGGGAGAT GTTCCGGAG 3360
 CAGAAGACGT CTCCCTTCT CTGCTCACC TGGTCGCCAA TCCATGCTCT CTTTCTTTTC 3420
 TCTGTCTACT CCTTATCCCT TGGTTTAGAG GAACCCAGA TGTGGCCTT AGCAAACTG 3480
 10 GACATATGCC AAACCCACTC ATGACTGCAT GACGGAGCCG AGCCATGTT CTTTACACT 3540
 CGCTGTGTC ACATCTCAGG GAACCTGACC TCAGGCACAC CTGCGAGAAG GCAAGGCCCT 3600
 GCCCTGCCCA ACCTCTGTGG TCACCCATGC ATCTTCCACT GGAACGTTT ACTGCAACA 3660
 CACCTTGGAG AAGTGGCATC AGTCAACAGA GAGGGGAGG GAAGGAGACA CCAAGCTCAC 3720
 CCTTCTCAT GACCGAGGT TCCCACTCTG GGCAGGCC CTCACTGCA AAGGATTTGT 3780
 15 AGATAACAT GACTGTGTTG TTTTAACCAA TAACAGCTT CTTATAATGA TTTTCTTACT 3840
 AATGATACT ACAAGTTTCT AGCTCTCACA GACATATAGA ATAAGGGTT TTGCATAATA 3900
 AGCAGTTGT TATTAGGTT AACAATATTA ATTCAGGTT TTAGTTGGA AAAACAATTC 3960
 CTGTAACTCT CTATTCTCTA TAATTGTAGT AATTGCTCTA CAGATAATGT CTATATATTG 4020
 GCCAACTGG TGCATGACA GTACTGTATT TTTTATACC TAAATAAGA AAAATCTTTA 4080
 20 GCCTGGGCAA CAAAAAA

Seq ID No: 105 Protein sequence:
 Protein Accession #: NP_001786.1

25 1 11 21 31 41 51
 | | | | | |
 MQRMLMLLAT SGACLGLLAV AAVAAGANP AQRDTHSLLP THRRQKRDWI WNMQHIDBSEK 60
 NTSLEPHRVGK IKSSVSRKNA KYLLKGEYVG KVFVRDAETG DVFAERLDR ENISEYHLTA 120
 30 PTVDKDTGEN LETPSSPTIK VHDVNDNWPV FTHRLPMASV PESSAVGTSV ISVTAVDADD 180
 FVVDGSHAVM YQILKGKEYF AIDNSGRIIT ITKSLDRKQ ARYHIVVEAR DAQGLRGDSG 240
 DTATLVLTQD INDNPFPTTQ TKYTFVVPED TRVGTSGVSL FVEDPDEPON RMTKYSILRG 300
 DYQDAFTIET NPAHNEGIIK PMKPLDYEYI QQYSFIVREAT DPTIDLRYMS PPAGNRAQVI 360
 INTKGDLYNE IPQPPYHFQ LKRNQKKPLI GTVLAMPDPA ARHSIGYSIR RTSKDGQFPR 420
 35 KPYQPKVCEN AVHQQLVLQI SAIDKDIPTP NVKFKFTLMT ENNFTLDNH DNTANITVKY 540
 GGFPRREHTKV HFLFVVISDN GMPSRGTST LTVAVCKCNE QGEFTFCEDM AAQGVVSIQA 600
 VVAILLICIT ITVTLLIFL RRLRKQARA HGKSVPEIHE QLVTYDEBGG GEMDTTSYDV 660
 SVLNSVRERG AKPPRPALDA RPSLYAQVQK PPRHAPGAHG GPGEAAMIE VKKDRADHDG 720
 40 DFPFYDTLHI YGEGSESLA ESLSSLTDS SDSVDYDPL NDWGRPFKML AELYGSDPRE 780
 ELLY

Seq ID No: 106 DNA sequence
 Nucleic Acid Accession #: none found
 Coding sequence: 1-474 (underlined sequences correspond to start and stop codons)

45 1 11 21 31 41 51
 | | | | | |
ACAGTACTCT GTGCAAAAAA CCTGGTGAAA AAGGATTTTT TCCGACTTCC TGATCCATTT 60
 GCTAAGSTGG TGGTTGATGG ATCTGGGCAA TGCCATTCTA CAGATACCTG GAAGAATAAG 120
 50 CTTGATCCAA AGTGAATCA GCATTATGAC CTGTATATTG GAAAGTCTGA TTCACTTACG 180
 ATCAGTGTAT GGAATCAACA GAAGATCCAT AAGAACAAG GTGCTGGATT TCTCGTGTG 240
 GTTCGTCTTC TTTCCAATGC CATCAACGCG CTCAAAGACA CTGGTTATCA GAGGTGGAT 300
 TTATGCAAAAC TCGGGCCAAA TGCAATGAT ACAGTTAGAG GACAGATAGT AGTAAGTCTT 360
 55 CAGTCCAGAG ACCGGAATAG CACAGGAGGA CAAGTTGTGG ACTGCAGTCG TTTATTGAT 420
 AACGATTTAC CAGACGGAGC TCATTATTG TGGACTTGGG AAGATAGATG TTAATGACTG 480
 GAAGTAAAC ACCCGTTAA AACACTGTAC ACCAGACAGC AACATTGTCA AATGGTCTG 540
 GAAAGCTGTG GAGTTTITG ATGAAGAGCG ACCAGCAGA TTGCTTCACT TGTGACAGG 600
 ATCCTCTCGA GTGCTCTGCG AGGGCTTCAA AGCATTCGAA GGTGCTGCGG GCCCGAGACT 660
 60 CTTTACCATA CACCAGATTG ATGCTGACAC TAACAACCTG CCGAAAGCCC ACCTTGCTT 720
 CAATCGAATA GACATTCCAC CCTATGAAAG CTATGAAAG CTATATGAAA AGCTGCTAAC 780
 AGCCATTGAA GAAACATGTG GATTTCGTGT GGAATGACAA GCTTCAAGGA TTTACCCAGG 840
 AC

65 Seq ID No: 107 Protein sequence:
 Protein Accession #: none found

70 1 11 21 31 41 51
 | | | | | |
 TVLCAKMLVK KDFRLPDPF AKVVVDGSGQ CHSTDTVKNT LDPKWNQHYD LYIGKSDSVT 60
 ISVWNHKKIH KQAGAGPLGC VRLLSNAINR LKDTGYQRLD LCKLGPNDND TVRGQIVVSL 120
 QSRDRIGTGG QVVDCHRLPD NDLPDGAHYL WTWKRC

75 Seq ID No: 108 DNA sequence
 Nucleic Acid Accession #: NM_002318.1
 Coding sequence: 248-2572 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
5  ACTCCAGCGC GCGGCTACCT ACGCTTGGTG CTGCTTTCT CCAGCCATCG GAGACAGAG 60
   CGGCCCCCTC TGCTCGAGAA AGGGGCTCAG CGCGGCGGGA AGCGGAGGGG GACCACCGTG 120
   GAGAGCGCGG TCCCAGCCCG GCCACTGCGG ATCCCTGAAA CCAAAAAGCT CCTGCTGCTT 180
   CTGTACCCCG CCTGTCCCTC CCAGCTGCGC AGGGCCCTTT CGTGGGATCA TCAGCCCGAA 240
   GACAGGGATG GAGAGGCTCT TGCTGCTCCA CCTCTGACGC TGCTGGCTCA TGCTGGCCCT 300
10  CCTGTCCCCC CTGAGCCTGG CACAGTATGA CAGCTGGCCC CATTACCCCG AGTACTTCCA 360
   GCAACCGGCT CCTGAGTATC ACCAGCCCCA GGGCCCGGCC AACGTGGCCA AGATTGAGCT 420
   GCGCTTGGCT GGGCAGAGAA GGAAGCACAG GAGGGGCGGG GTGGAGGTGT ACTATGATGG 480
   CCAGTGGGGC ACCGTGTGGG ATGACGACTT CTCCATCCAC GCTGCCACOG TCGTCTGCGG 540
   CGAGCTGGGC TATGTGGAGG CCAAGTCCCT GACTGCGACG TCCTCCTACG GCAAGGGAGA 600
15  AGGCCCCATC TGGTTAGACA ATCTCCACTG TACTGGCAAC GAGGCGACCC TTGCGAGCATG 660
   CACCTCCAAT GGCTGGGCGG TCACTGACTG CAGGACACAG GAGGATGTGG GTGTGGTGTG 720
   CAGCGACAAA AGGATTCCTG GGTTCAAATT TGACAATTGG TTGATCAACC AGATAGAGAA 780
   CCTGAATATC CAGGTGGAGG ACATTCCGAT TCGAGCCATC CTCTCAACCT ACCGCAAGCG 840
   CACCCCGAGT ATGGAGGGCT ACGTGGAGGT GAAGGAGGGC AAGACCTGGA AGCAGATCTG 900
20  TGACAAGCAC TGGACGCCCA AGAATTTCCG CGTGGTCTGC GGCATGTTTG GCTTCCCTGG 960
   GGAGAGGACA TACAATACCA AAGTGTACAA AATGTTGGCC TCACGAGGGA AGCAGGCTA 1020
   CTGGCCATTC TCCATGGACT GCACCGGCAC AGAGGCCAC ATCTCCAGCT GCAAGCTGGG 1080
   CCCCAGAGTG TCACTGGACC CCATGAAGAA TGTCACTGCG GAGAATGGGC TGCGGCGGT 1140
   CGTGAATTGT GTGCTGGGCG AGGTCTTCAG CCTTGACGGA CCCTCGAGAT TCGGAAAGC 1200
25  ATACAAGCCA GAGCAACCCC TGGTGGGACT GAGAGGCGGT GCCTACATCG GGGAGGGCGG 1260
   CTGGGAGGTG CTCAAAAATG GAGAATGGGG GACCGTCTGC GACGACAAAT GGGACCTGGT 1320
   GTGCGGAGT GTGGTCTGCA GAGAGCTGGG CTTTGGGAGT GCCAAGAGG CAGTCACTGG 1380
   CTCCCGACTG GGGCAAGGGA TCGGACCCAT CCACCTCAAC GAGATCCAGT GCACAGGCAA 1440
   TGAGATGCTC ATTATAGACT GCAAGTTCAA TGCGAGTCTC CAGGGCTGCA ACCACAGGGA 1500
30  GATGCTGGT GTGAGATGCA ACACCCCTGC CATGGGCTTG CAGAAGAAGC TGCGCTGAA 1560
   CGCGCGCGCG AATCCCTACG AGGGCCGAGT GGAGGTGCTG GTGGAGAGAA ACGGCTCCCT 1620
   TGTGTGGGGG ATGGTGTGTG GCCAAAAC TG GGCATCGTG GAGGCCATGG TGGTCTGCGG 1680
   CCAGCTGGGC CTGGGATTCG CCAGCAAGCG CTTCAGAGAG ACCTGTGATT GGCACGAGGA 1740
   TGTCAACAGC AACAAAGTGG TCATGAGTGG AGTGAAGTGC TCGGGAACGG AGCTGTCCCT 1800
35  CGCCACTGTC CGCCACGAGC GGGAGGACGT GGCCTGCCCC CAGGGCGGAG TGCATACGG 1860
   GGCCGAGGTT GCCTGCTCAG AAACCGCCCC TGACTGGTTC CTCAATGGGG AGATGTGCA 1920
   GCAGACCAAC TACCTGGAGG ACCGCGCCAT GTTCATGCTG CAGTGTGCCA TGGAGGAGAA 1980
   CTGCTCTGCG GCCTCAGCGG CGCAGACCGA CCCCACCAAG GGCATCGGCC GGCCTCTGCG 2040
   CTTCTCTCCC CAGATCCACA ACAATGGCCA GTCCGACTTC CGGCCCAAGA ACGGCGGCA 2100
40  CGCTGGATG TGGAACGAGT GTCCACAGCA CTACCAAGC ATGAGAGTGT TCAACCACTA 2160
   TGACCTGTGC AACCTCAATG GCACCAAGGT GGCAGAGGGC CACAAGGCCA GCTTCTGCTT 2220
   GGAGGACACA GAATGTGAAG GAGACATCCA GAAGAATTAC GAGTGTGCCA ACTTCCGCGA 2280
   TCAGGCGATC ACCATGGGCT GCTGGGACAT GTACCGCCAT GACATCGACT GCCAGTGGGT 2340
   TGACATCACT GACGTGCCCC CTGGAGACTA CTTGTTCCAG GTTGTATTAT ACCCCAATT 2400
45  CGAGGTGACA GAATCCGATT ACTCCAACAA CATCATGAAA TGCAAGAGCC GCTATGACGG 2460
   CAGACTTGGT TGGATGTACA ACTGCCACAT AGGTGGTTCC TTCAAGCAAG AGAAGGAAA 2520
   AAGTGTGAG CACTTCAGCG GGCTCTTAAA CAACAGCTG TCCCGCAGT AAGAGAGCT 2580
   GCGTGTGCAA CTCCTGTCTT CAGGCCACAC CACATCTTCC ATGGGACTTC CCCCACAA 2640
   CTGAGTGTGA AOGAATGCCA CGTGCCTCA CCGACGCGGG CCCCACCTCT GTCCAGACCC 2700
50  CTACAGCTGT GTCTAAGCTC AGGAGGAAAG GGACCCCTCC ATCATTCATG GGGGCTGCT 2760
   ACCTGACCTT TGGGCGCTGA GAAGGCCCTT GGGGGGTGGG GTTTGTCCAC AGAGCTGCTG 2820
   GAGCAGCACC AAGAGCCAGT CTTGACCGGG ATGAGGCCCA CAGACAGGTT GTCATCAGCT 2880
   TGTCCCATTC AAGCCACCGA GCTCACCACA GACACAGTGG AGCCGCGCTC TTCTCCAGTG 2940
   ACACGTGGAC AATGCGGGC TCATCAGCCC CCCCAGAGAG GGTGAGGCGG AACCCCATTT 3000
55  CTCTCTCTCT TAGGTCAATT TCAGCAAACT TGAATATCTA GACCTCTCTT CCAATGAAC 3060
   CCTCCAGTCT ATTATAGTCA CATAGATAAT GTTGCCACGT GTTPTCTGAT TTGGTGAAGCT 3120
   CAGACTTGGT GCTTCCCTCT CCACAAACCC CACCCCTTGT TTTTCAAGAT ACTATTATTA 3180
   TATTTTCACA GACTTTTGAA GCACAAATTT ATTGGCATTT AATATTGGAC ATCTGGGCCC 3240
   TTGGAAGTAC AATCTAAGG AAAAACCAAC CCACTGTGTA AGTGACTCAT CTTCCTGTTG 3300
60  TTCCAATTCT GTGGGTTTTT GATTCAACGG TGCTATAACC AGGGTCTGCG GTGACAGGGC 3360
   GCTCACTGAG CACCATGTGT CATCAGACAC ACTTACACAT ACTTGAACCT TGAATAAAA 3420
   GAAAGATTTA TG

```

Seq ID No: 109 Protein Sequence
 Protein Accession #: NP_002309.1

```

1      11      21      31      41      51
70 MERPLCSHLC SCLAMLALLS PLSLAQYDSW PHYPEYFQQP APEYHQFQAP ANVAKIQLRL 60
   AGQKRKHSEG RVEVYDGGW GTVDDDFSI HAAHVVCREL GYVEAKSWTA SSSYKGEGP 120
   IWLNLHCTG NEATLAACTG NGWGVTDCKH TEDVGVVCSG KRIPGFKPDN SLINQIENLN 180
   IQVEDIRLRA ILSTYRKRTF VMEGVVEVKE GKIWKQICDK HWTAKNSRVV CQMFGPFGER 240
   TYNTKVYKMF ASRRKQRYWF PSMDCTGTEA HISSCLGPQ VSLDPMKIVT CENGLPAVVS 300
75 CVPQOVYSPD GPSRFKAYK PEQFLVLRGQ GAYIGGRVE VLKNGEWGTV CDRKDLVSA 360
   SVVCRLGFG SAKRAVTGSR LGQGIGPIHL NBIQCTGNEK SIIDCKPNAE SQGCNHBDA 420

```

5 GVRCTPAMG LQKKLRINGG RNPYBGRVEV LVERNGSLVW GMVCGQNWGI VEMMVVCRQL 480
 GLGFASNAFQ BTWYHGDVW SNKVMGKVK CSGTELSLAH CRHGDGVAC PQGGVQYAG 540
 VACSTAPDL VLAEMVQQT TYLEDPMFM LQCAMBENCL SASAAQTDPT TGYRLLRFS 600
 SQIHNNQSD FRPKNGRHAW IWHDCRHYH SMEVPTHYDL LNLNGTRVAB GHKASFCLD 660
 TECEGDIQKN YECANFGDQG ITMGCDMYR HDIDCQWVDI TDVPPGDYLF QVINPNEFV 720
 AEDSYNNIN KCRSRYDGRH IWMYNCHIGG SPSEBTEKKF BHFSGLLANQ LSPQ

Seq ID NO: 110 DNA sequence

Nucleic Acid Accession #: none found, CAT_73007_3

10 Coding sequence: 1-495 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 15 QGGACGGTGG GGTGACCCCA CGCTCCGCC CACGGTCCG TATGACAGA GCCTCCACTG 60
 GCTGCTGCCT GCGCGCCACA TACCCAGCTG ACATGGGCAC CGCAGGAGCC ATGCAGCTGT 120
 TCGGGTGATC CTGGGCTTCC TCCTGTTCGG AGGCCACAAC TCCCAGCCCA CAATGACCCA 180
 ACCTCTAGCT CTCAGGAGG CCTTGGGGGT CTAAGTCTGA CCACAGAGCC AGTTTCTTCC 240
 ACCAGGAGTA CATCCCTTCC TCAGAGGCTA ACAGGCCAAG CCATCTGTCC AGCACTGGTA 300
 CCCAGGGCGA GGTGTCCCCA GCAGTGGAAG AGACGGAGGC ACAAGCAGAG ACACATTTC 360
 20 ACTGTTCGCC CCATTCACAC CACCATGAGC CTGAGCATGA GGAAGATGC GACCATCTCT 420
 CAGCCCCAC GTCAAGACT GTGCTCACTG TGGCTGCAIT TGGATGAGG TCGGGTGGAG 480
 GCCCACTCTG GCTAGGGGGC GGCAGGCTGA GAGCTCACCT GTTCAGCAGA GAAGTGGAAC 540
 CACTTTGCTC CTGAGCCTG TCTACCACAG TGTATCAGC TTCATTGTCA TCCTGGTGGT 600
 25 TGTGTATCA TCTAGTTGG TGTGTCAGC CTGAGGGTTC AGTGTGGAA GAGCAAGGAG 660
 TCTGAAGATC CCAGAACCTG GGAGTACAGG GCGTGTCTGA CAAGCTGTCT ACAGACCATG 720
 GCGAGAACGA CAGCATCGCC CATTATCACA TGGAGACAT CACACGACT AGGGCAACAC 780
 GCACCTAGCA GCGAGCATCA AAGGAGCCTA CGCATGGCCC AGACTGAGAG CAAGCACAAA 840
 GGGC

30 Seq ID NO: 111 Protein sequence:

Protein Accession #: none found, CAT_73007_3

35 1 11 21 31 41 51
 | | | | |
 RTRGSTHASA HASVWTEPPL AAACPHTQL TNAPQEPSC LGDPGLFPVP RPQLPAHNDP 60
 TSSSGGGLGG LSLTTEPVSS TQDTSLPQRL TGQAICPALV PRRCPPQWK RRRHKQRHIS 120
 TVPPNSTIMS LSMREDATIL PAPRQLCSL WLHLGWSRVE AHSQ

40 Seq ID NO: 112 DNA sequence

Nucleic Acid Accession #: NM_005424.1

Coding sequence: 37-3453 (underlined sequences correspond to start and stop codons)

45 1 11 21 31 41 51
 | | | | |
 CGCTCGTCTT GGCTGGCCTG GGTGGGCTC TGGAGTATGG TCTGGCGGGT GCCCCCTTTC 60
 TTGCTCCCCA TCTCTTCTT GGCTTCTCAT GTGGGCGCGG CGGTGGACCT GACGCTGCTG 120
 50 GCCAACCTGC GGTCTACGGA CCCCAGCGC TTCTTCTCTA CTGCTGTGTC TGGGAGGSCC 180
 GGGGCGGGGA GGGGCTCGGA CGCTGGGGC CGGCCCTGTC TGTGGAGAA GGACGACCGT 240
 ATCGTGGGCA CCGCGCCCGG GCCACCCCTG CGCTGGGCGC GCAACGGTTC GCACCAAGTC 300
 AGCTTGGGCG GCTTCTCCAA GCCCTGGGAC CTGCTGGGCG TCTTCTCTG CGTGGGCGGT 360
 GCTGGGCGCG GCGCACGCG CGTCTATPAC GTGCACAACA CGCCTGGAGC CCACCTGCTT 420
 CCAGACAGG TCACACACAC TGTGAACAAA GGTGACACCG CTGTACTTTC TGACGCTGTG 480
 55 CACAGGAGAG AGCAGACAGA CGTGATCTGG AAGAGCAACG GATCCTACTT CTACACCCTG 540
 GACTGGCATG AAGCCCAGGA TGGGCGGTTT CTGCTGCAGC TCCCAATGT GCAGCCACCA 600
 TCGAGCGGCA TCTACAGTGC CACTTACCTG GAAGCCAGCC CCTGGGCG CGCTTCTTTT 660
 CGGCTCATCT TCGGGGTTG TGGGGCTGGG CGCTGGGGCG CAGGCTGTAC CAAGGAGTGC 720
 CCAGGTTGCC TACATGGAGG TGTCTGCCAC GACCATGACG GCGAATGTGT ATGCCCCCTT 780
 60 GGCTTCACTG GCACCCGCTG TGAACAGGCC TGCAAGAGAG GCGTTTGTGG GCAGAGCTGC 840
 CAGGAGCAGT GCCCAGGCAT ATCAGGCTGC CGGGGCTCA CCTTCTGCTT CCCAGACCCC 900
 TATGGCTGTG CTTGTGGATC TGGCTGGAGA GGAAGCCAGT GCGAAGAGC TTTGTGCCCT 960
 GGTCAATTTT GGGCTGATTG CCGACTCCAG TGCCAGTGTG AGAATGGTGG CACTTGTGAC 1020
 65 CGGTTCAGTG GTTGTGCTG CCCCTCTGGG TGGCATGGAG TGCACTGTGA GAAGTCAGAC 1080
 CGGATCCCCC AGATCTCAA CATGCCCTCA GAATGGAGT TCACTTAGA GACGATGCCC 1140
 CGGATCAACT GTGACGTGC AGGGAACCCC TTCCCGTGC GGGCAGCAT AGAGCTACGC 1200
 AAGCCAGAGC GCACTGTGCT CCGTGTCCAC AAGGCCATTG TGGAGCCAGA GAAGACCACA 1260
 GCTGAGTTTG AGTGCCCCG CTTGTGTTCT GCGGACAGTG GGTCTGGGA GTGCGGTGTG 1320
 70 TCCACATCTG GCGGCCAGA CAGCCGGCGC TTCAGGTCA ATGTGAAAGT GCCCCCGTGT 1380
 CCCCTGGCTG CACTCGGCT CCGACCAAG CAGAGCCGCC AGCTGTGGT CTCCCGCTGT 1440
 GTCTGTTCTG CTGGGATGG ACCCATCTCC ACTGTCCGCC TGCACTACCG GCCCCAGGAC 1500
 GAGACCAATG ACTGTCGAC CATTGTGGT GACCCAGTG AGACGCTGAC GTTAATGAAC 1560
 CTGAGGCCAA AGACAGGATA CAGTGTTCGT GTGCAGCTGA GCGGCCAGG GGAAGGAGGA 1620
 75 GAGGGGCTCT GGGGCTCTCC CACCCTCATG ACCACAGACT GTCTGAGCC TTTGTTCAG 1680
 CCGTGTGTGG AGGGCTGGCA TGTGAAAGGC ACTGACCGG TGCGAGTGTG CTGCTCCTTG 1740
 CCTTGTGTGC CCGGGCCACT GGTGGGCGAC GGTTCCTGTC TGCGCTGTG GACGCGGACA 1800

5
 10
 15
 20
 25
 30
 35

```

CGGGGGCAGG AGCGGGGGGA GAACGCTCTCA TCCCCCAGG CCCGCACTGC CCTCCTGACG 1860
GGACTCACGC CTGGCACCCA CTACCAGCTG GATGTGCAGC TCTACCACTG CACCCCTCTG 1920
GGCCCGGCTC CGCCCGCTGC ACACGTGCTT CTGCCCCCA GTGGGCTCC AGCCCCCGA 1980
CACCTCCACG CCCAGGCCCT CTCAGACTCC GAGATCCAGC TGACATGGAA GCACCGGAG 2040
GCTCTCCCTG GSCAATATC CAAGTACGTT GTGGAGGTGC AGGTGGCTGG GGTGACAGGA 2100
GAACCACTGT GGATAGACGT GGACAGGCTT GAGGAGACAA GCACCATCAT CCGTGGCCTC 2160
AACGCCAGCA CGCGTACCT CTTCCGCTG CGGCCAGCA TTCAGGGGCT CGGGGACTGG 2220
AGCAACACAG TAGAAGAGTC CACCTGGGC AACGGGCTGC AGGCTGAGGG CCGAGTCCAA 2280
GAGAGCGGGG CAGCTGAAGA GGGCTGGAT CAGCAGCTGA TCTGGCGGT GGTGGGCTCC 2340
GTGTCTGCCA CCGCTCTCAC CATCTGGCC GCCTTTTAA CCTGGTGTG CATCCGAGA 2400
AGCTGCTCTC ATCGAGAGC CACCTTCACC TACCAGTCAG GCTCGGGCGA GAGAGCCATC 2460
CTGCAGTTCA GCTCAGGAC CTTGACACTT ACCCGCGGCG CAAACTGCA GCCCGAGCCC 2520
CTGGAGGTTA CAGTCTAGA GTGGAGGAC ATCACCCTTG AGGACCTCAT CGGGAGGGG 2580
AACTTGGGCG AGTTCATCG GGCCATGATC AAGAAGAGAC GGCTGAAGAT GAACGAGCC 2640
TCTAATAATC TGAAGAAGTA TGCTCTGAA AATGACCATC GTGACTTTC GGGAGAACTG 2700
GAAGTCTCTG GCAAATTGGG GCATCAACCC AACATCATCA ACCTCTGGG GGCTGTAA 2760
AACCGAGGTT ACTTGTATAT CGCTATTGAA TATGCCCTT ACGGGAACCT GCTAGATTTT 2820
CTGGGAGAAA CGCGGCTCTT AGAGACTGAC CCGACTTTTG CTCGAGAGCA TGGGAGGCC 2880
TCTAAGCTTA GCTCCGGCA GCTGCTGGT TCGCCAGTG ATGCGGCCAA TGGCATGAG 2940
TACCTCAGTG AGAAGCAGTT CATCCACAGG GACCTGGCTG CCGCGAATGT GCTGGTCGGA 3000
GAGAAGCTAG CCTCCAGAT TGCAGACTTC GGCCTTCTC GGGGAGAGGA GGTATTATGT 3060
AAGAAGACGA TGGGGCGTCT CCTGTGGCG TGATGGCCA TTGAGTCCCT GAACACAGT 3120
GTCTATACCA CCAAGAGTGA TGTCTGTGCC TTGGAGTCC TTCTTTGGGA GATAGTGAGC 3180
CTGGAGGTTA CACCTACTG TGGCATGACC TGTCCGAGC TCTATGAAA GCTGCCCCAG 3240
GGCTACCGCA TGGAGCAGCC TCGAAACTGT GAGCATGAAG TGTAAGAGCT GATGCTCAG 3300
TGCTGGCGGG ACCGTCCTTA TGAGCGAACC CCCTTTGCC AGATTGCGCT ACAGTAGGC 3360
CGCATGCTGG AAGCCAGGAA GGCCTATGT AACATGTGCG TGTATGAAA CTTCACITAC 3420
GGCGGCTGTT ATGCCACAGC TGAGGAGGCC TGGCTGCCA TCCAGCCAGA ACGTGGCTCT 3480
GCTGGCGGGA GCAAACTCTG CTGTCTAACC TGTGACCACT CTGACCTTCA CAGCTCTGA 3540
30 CTTAAGCTTA CTCAAGGAT TTTTCTAACC TAAGGGAGAA AAAAGGGAT CTGGGATGG 3600
GGTGGCTTCA GGGGAAGTGG GTTCCCATGC TTTGTAGGT TCTCATAGCT ATCCTGGGCA 3660
TCTTCTTTC TAGTTCAGCT GCCCCACAGG TGTGTTTCCC ATCCCACTGC TCCCCAACCA 3720
CAAAACCCCA CTCAGCTCC TTGCTTAAG CCAGCACTCA CACCACTAAC ATGCCCTGTT 3780
CAGTACTCTC CACTCCCGGC CTGTCTTCA GAAAAAATA AATGTTCTAA TAAGCTCCAA 3840
AAAAA
  
```

Seq ID No: 113 Protein sequence
 Protein Accession #: NP_005415.1

40
 45
 50
 55
 60

```

1 11 21 31 41 51
MWRVPPFFLL PFLFLASHVG AAVDLTLIAN LRLTDPQRF LTVSGEAGA GRGSDAWOPP 60
LLLEKDDRIY RTPPPPLRL ARNOSHQVTL RGFSKPSDLV GVFSVCGAG ARRTVIYVH 120
NSPGAHLPLD KVTHTVNKGD TAVLSARVHK EKQTDVINKS NGSYFYTLGW HEAQDGRFLL 180
DLGFNVQPPSS GIYSATYLEA SPFGSAFFRL IVRGCGAGRW GPGCTKECPG CLHGGVCHDH 240
QGECVCPGPG TGTCEQACR BGRFGQSCQE QCPGISGCRG LTFCLDPDYG C8CGSGWRGS 300
QCQACAPGH FGADCRQCQ CQNGGTCDRF GGCVCPSGWH GVHCEKSDRI PQILNMASEL 360
EFNLETMPRI NCAAGNPPF VRSIELRKP DGTVLLSTKA IVEPEKTAE FVPRVLVAD 420
SGFWEKRVST SGGQDSRRFK VNVKVPVPL AAPRLITKQS RQLVVSPLVS FSGDGPISITV 480
RLHYRFPDST MDWSTIVVDF SENVTLMNLR PKTGYSVRVQ LSRPGEGEGE AWGPPTLMTT 540
DCPEPLLPFW LBGWHVETD RLRVSNLPL VPGPLVGDGF LLRLWDGTRG QERRENVSSP 600
QARTALLTGL TPQTHVQLDV QLYHCTLLGP ASPPAHVLLP PSGPPAPRHL HAQALSDSEI 660
QLTWKHPHAL PGPISKYVVE VQVAGGAGDP LWIDVDRPEE TSTIIRGLNA STRYLFMRRA 720
SIQGLGDWSN TVEESTLGMG LQAGFPVQES RAAEBGLDQ LILAVVGSVS ATCLTILAA 780
LTLVCIRRS C LHRRTFTYQ SGSGRETIQ FSSGTLTLTR RPKLQPEPLS YPVLWEEDIT 840
FEDLIGBNF GQVIRAMIK DGLKMNAAIK MLKEYASEND HRDPAGELEV LCKLGHHPNI 900
INLLGACKNR GYLIAIEYA PYGNLDFLR KSRVLETPA FAREHGTAST LSSRQLLRPA 960
SDAANGMYL SEKQPIHDL AARNVLVGN LASKIADPGL SRGEVYVKK TMRPLVVRM 1020
AIESLAYSIV TTKSDVVSFG VLLWEIVSLG GTPYCGMTCA ELYEKLPQGY RMEQPRNCDD 1080
60 EVVELARQCN RDRPYRPPF AQIALQLGRM LEARKAYVM SLFENFTYAG IDATABEA
  
```

Seq ID NO: 114 DNA sequence

Nucleic Acid Accession #: NM_002632.1

65 Coding sequence: 322-771 (underlined sequences correspond to start and stop codons)

70
 75

```

1 11 21 31 41 51
GGGATTCGGG CGGCCAGCT ACGGGAGGAC CTGGAGTGGC ACTGGCGGCC CGACGGACCA 60
TCCCGGGGAC CGCCTGCCC CTGGCGGCCC CGCCCGCCCG GCGCGCTCCC CGTCGGGTTT 120
CCAGCCACA GCCTTACCTA CGGCTCCTCG ACTCGCAAG GCTTCCAGAA GATGCTGAA 180
CCACCGGCCG GGGCCTCGGG GCAGCAGTGA GGGAGGCGTC CAGCCCCCA CTCAGCTCTT 240
CTCCTCCTGT GCCAGGGGCT CCGCGGGGGA TGAGCATGGT GGTTTTCCCT CGGAGCCCCC 300
TGCTCTGAGC GTCTGAGAA GATGCGGCTC ATGAGGCTGT TCCCTTGCCT CTCGAGCTC 360
75 CTGCGCGGCC TGGCGCTGCC TGCTGTGCC CCGCAGCAGT GGGCCTTGTG TGCTGGGAC 420
GGCTCTGTAG AGGTGGAAGT GGTACCCCTT CAGGAAGTGT GGGCGCCGAG CTACTGCCGG 480
  
```

5 GCGCTGGAGA GGCTGGTGA CGTCGTGTCC GAGTACCCCA GCGAGGTGGA GCACATGTTT 540
 AGCCCATCTT GTGTCTCCCT GCTGCGCTGC ACCGGCTGCT GCGGCGATGA GAATCTGCAC 600
 TGTGTGCGCG TGGAGACGGC CAATGTCAACC ATGCAGCTCC TAAAGATCCG TTCTGGGGAC 660
 CGGGCTCTCT ACGTGGAGCT GAGCTTCTCT CAGCAGGTTT GCTGCGAATG COGGCTCTGT 720
 CGGGAGAGA TGAAGCCGGA AAGGTGCGGC GATGCTGTTC CCGGAGGTA ACCACCCCTT 780
 TGGAGGAGAG AGACCCCGCA CCGGCTCGT GTATTATTA CCGTCAACT CTTCAGTGAC 840
 TCCTGCTGGT ACCTGCCCTC TATTATATAG CCAACTGTTT CCTGTCTGAA TGCTCGCTC 900
 CCTTCAGAC GAGGGGAGG GAAGACAGG ACCCTCAGGA ATTCAGTGCC TTCAACAACG 960
 10 TGAGAGAAAG AGAGAAGCCA GCCACAGACC CCTGGAGCT TCCGCTTGA AAGAAGCAAG 1020
 ACACGTGGCC TCGTAGGGG CAAGCTAGGC CCCAGAAGCC CTGAGAGTCT CCAGGGGCTT 1080
 GCAGAAGGAA AGAAGGGGGC CCTGCTACCT GTTCTTGGG CTCAGGCTCT GCACAGACAA 1140
 GCAGCCCTTG CTTTCGGAGC TCCTGTCCAA AGTAGGGATG CGGATTCCTG TGGGGCCGCC 1200
 AGGCTCTGGT GGTGGGAAGG CCGGAGCGG GGGGAGGGA TTCAGCCACT TCCCTCTCTT 1260
 15 CTTCTGAGA TCAGAACATT CAGCTCTGGA GAACAGTGTG TGCTGGGGG CTTTGGCCAC 1320
 CCTTGTGCC CCGTATCTC CCCTCACACT TTGCCATTG CTTGTACTGG GACATTGTTT 1380
 TTTCCGGCG AGGTGCCACC ACCCTGCCCC CACTAAGAGA CACATACAGA GTGGGCCCG 1440
 GGCTCTGAG AGAGCTGCCT GGATGAGAAA CAGCTCAGCC AGTGGGGATG AGGTCAACAG 1500
 GGGAGAGCG TGTGCTGCC AGCTGAAGGC AGTGGCAGGG GAGCAGGTTT CCCAGGGCC 1560
 20 CTGACACCC CACAAGCTGT CCTGCAGGG CCATCTGACT GCCAAGCCAG ATTCTCTTGA 1620
 ATAAAGTATT CTAGTGTGA AACGC

Seq ID No: 115 Protein sequence
 Protein Accession #: NP_002623.1

25 1 11 21 31 41 51
 MPVMRLFPFC LQLLAGLALP AVPPQWALS AGNGSSEVEV VPFQEVWGRS YCRALERLVD 60
 VVSEYPSEVE HMFSPSCVSL LRCTGCCGDE NLHCVPVSTA NVTMLLKLIR SGRDPSYVEL 120
 30 TFSQHVRCRC RPLREKMKPE RCGDAVPRR

Seq ID NO: 116 DNA sequence
 Nucleic Acid Accession #: NM_007361.1
 Coding sequence: 1-4131 (underlined sequences correspond to start and stop codons)

35 1 11 21 31 41 51
 ATGAGGGGG ACCGGGTGGC CGGGGGCGCG GTGCTGTGCT CGTTACCACT GCTACTGCTG 60
 CTGCACTTGC TAATGTTGCG GCGCGCGCGG CTGCACCCAG ACGAGCTCTT CCCACACGGG 120
 GAGTCTGGGT GGGACCACTT CCTGCAGGAA GCGGACGACG TAAAGCTCAG CCGTGGTGAA 180
 40 GCTGGGAAAT CCCCTGCACT TCTTACGAAG CCGGATTGAG CAACCTCTAC GTGGGCACCA 240
 CAGGATCAT CTCCACTCAG GACTTCCCA GGGAAACGCA GTATGTGAC TATGATTTC 300
 CCACCGACTT CCGGCGCATC GCGCCCTTTT TGGCGGACAT CGACACGAGC CACGGCAGAG 360
 GCGAGTCTCT GTACCGAGAG GACACCTCCC CCGCAGTGCT GGGCCTGGCC GCCCGCTATG 420
 45 TGGCGCGCTG CTTCCCGCGC TCTGCGCGCT TTTTACCCCC ACCCAGCCTT TCTTGGCCAC 480
 CTGGAGGACG GTAGGCGCTT ACGAGGAGGT CAACCGCGGG CCGTGCCTCT GGGAGAGCTG 540
 AACATTTTCC AGGCAGTTTT GGCATCTGAT GGGTCTGATA GCTACGCCCT CTTTCTTTAT 600
 CAGTCCCAAG GCCTGCAATT CCTTGAAGCC GCGCCCAAG AGTCTTACAA TGTCCAGCTT 660
 CCTTCTCAG CTCGGGTGGG CTTCTGCCGA GGGGAGGCTG ATGATCTGAA GTCAGAAGGA 720
 50 CCATATTCCA GCTTGACTAG CACTGAACAG TCTGTGAAA ATCTCTATCA ACTAAGCAAC 780
 CTGGGGATCC CTGGAGTGTG GCGTTTCCAT ATCGGCAGCA CTTCCCGGTT GGACAATGTC 840
 AGGGCAGCTG CAGTGTGAGA CCTTCCGCTT GCCCACTCTT CTGTCCCTCT GGGAGGTTCC 900
 TTCAAGCATG CTACAGCCCT GGAAGTGAC TATAATGAG ACAATTGGA TTACTACGAT 960
 GTGAATGAGG AGGAAGCTGA ATACCTTCCG GGTGAACAG AGGAGGCATT GAATGGCCAC 1020
 55 AGCAGCATTG ATGTTTCCCT CCAATCCAAA GTGGATACAA AGCCTTTAGA GGAATCTTCC 1080
 ACCTTGGTAT CTCACACCAA AGAAGGAACA TCTCTGGGAG AGGTAGGGGG CCCAGATTTA 1140
 AAAGGCCAAG TTGAGCCCTG GGATGAGAGA GAGACCAGAA GCCCAGCTCC ACCAGAGGTA 1200
 GACAGAGATT CACTGGCTCC TTCTGGGAA ACCCCACCA CGTACCCCGA AAACGGGAAGC 1260
 ATCCAGCCCT ACCCAGATGG AGGGCCAGTG CCTTCGGAAA TGGATGTTCC CCCAGCTCAT 1320
 60 CCTGAAGAAG AATTTGTTCT TCGAAGTTAC CCTGCTTCAG GTCACACTAC ACCCTTAAAT 1380
 CGAGGGACGT ATGAGGTGGG ACTGGAAGAC AACATAGGTT CCAACACCGA GGTCTTCAAG 1440
 TATAATGCTG CCAACAAGGA AACCTGTGAA CACAACCA GACAATGCTC CCGCATGCC 1500
 TTCTGCAGCG ACTATGCCAC TGGCTTCTGC TGCCACTGCC AATCCAAGTT TTAATGGAAT 1560
 GGGAGCACT GTCTGCCTGA GGGGGCACCT CACGAGTGA ATGGGAAGT GAGTGGCCAC 1620
 65 CTCCAGCTGG GCCATACACC CGTGCACCTT ACTGATGTGG ACCTGCATGC GTATATCTGT 1680
 GGCATGATG GCAGAGCCTA CACGGCCATC AGCCACATCC CACAGCCAGC AGCCCAAGGC 1740
 CTCTCCCCC TCACACCAAT TGGAGGCTGT TTTGGCTGGC TCTTTGCTTT AGAAAAACCT 1800
 GGCTCTGAGA ACGGCTTCAG CCTCGCAGGT GCTGCTTTA CCGATGACAT GGAAGTTACA 1860
 TTCTACCCGG GAGAGGAGAC GGTTCGTATC ACTCAACTG CTGAGGAGCT TGACCCAGAG 1920
 70 AACTACCTGA GCATTAGAC CAACATTCAA GGGCAGGTGC CTTAGTCCC AGCAAAATTC 1980
 ACAGCCCACT TCTCTCCCTA CAAGGAGCTG TACCACTACT CCGACTCCAC TGTGACCTCT 2040
 ACAGTTTCCA GAGACTACT TCTGACTTTT GGTGCAATCA ACCAACATG GTCTTACCGC 2100
 ATCCACCAAG ACATCACTTA CCAGGTGTGC AGGCAGGCC CCAGACACCC GTCTTCCCTC 2160
 ACCACCCAG AGCTGAACGT GGACCGGGTC TTTGCTTGT ATAATGATGA AGAAGAGAGT 2220
 75 CTTAGATTG CTGTGACCAA TCAAAATTGG CCGGTCAAAG AAGATTGAGA CCCCACCTCG 2280
 GTGAATCCTT GCTATGATGG GAGCCACATG TGTGACACAA CAGCAGGCTG CCATCCAGGG 2340
 ACAGGTGTAG ATTACACCTG TGAGTGCSCA TCTGGGTACC AGGGAGATGG ACGGAATCTG 2400

5
10
15
20
25
30
35
40
45

GTGGATGAAA ATGAATGTGC AACTGGCTTT CATCGCTGTG GCCCAACTC TGTATGTATC 2460
AACTTGCTTG GAAGCTACAG GTGTGAGTGC CGGAGTGGTT ATGAGTTTGC AGATGACCGG 2520
CATACTTGCA TCTTGATCAC CCCACCTGCC AACCCCTGTG AGGATGGCAG TCATACCTGT 2580
GCTCTGCTGT GGCAGGCCCG GTGTGTTTAC CATGGAGGCA GCACGTTTCA CTGTGCTGTC 2640
CTGCTCGGTT ATGCGGGGGA TGGGCAACCAG TGCACCTGAT TAGATGAATG CTCAGAAAAAC 2700
AGATGTCAACC CTGCAGCTAC CTGTACAAAT ACTCCTGGTT CCTTCTCCTG CCGTTGTCAA 2760
CCCGGATATT ATGGGGATGG ATTTCAATGC ATACCTGACT CCACCTCAAG CCTGACACCC 2820
TGTGAACAAC AGCAGGCGCA TGGCCAGGCC CAGTATGCCT ACCCTGGGGC CCGTTCCAC 2880
ATCCCCAAT GCGACGAGCA GGGCAACTTC CTGCCCCAC AGTGTCTATG CAGCACTGGT 2940
TTCTGCTGGT GGTGTGACCC TGATGGTTCAT GAAGTTTCTG GTACCCAGAC TCCACCTGGC 3000
TCCACCCCGC CTCCTGTGAG ACCATCAGCA GAGCCCAACC AGAGGCCCCC GACCATCTGT 3060
GAGCGCTGGA GGGAAAAAAT GCTGGAGCAC TACGGTGGCA CCCCCGAGAG TGACCACTAC 3120
GTGCCCCAGT GCGATGACCT GGGCCACTTC ATCCCCCTGC AGTGCCACGG AAGAGCGGAC 3180
TTCTGCTGGT GTGTGGAACA AGATGGCAGA GAGGTGCAGG GCACCCGCTC CCAGCCAGGC 3240
ACCACCCCTG CGTGTATACC CACCGTCGCT CCACCCATGG TCCGGCCCCA GCCCCGGCCA 3300
GATGTGACCC CTCATCTGT GGGCACCTTC CTGCTCTATA CTCAGGGCCA GCAGATTGGC 3360
TACTTACCC TCAATGGCAC CAGGCTTCAG AAGGATGCAG CTAAGACCCCT GCTGTCTCTG 3420
CATGGCTCCA TAATGTTGGG AATTGATTAC GACTGCGGG AGAGGATGGT GACTGTGACA 3480
GATGTGCTG GACGACAAT CAGCGTGCC GGTCTGGAAC TGGGAGCAGA GCGTGAGACG 3540
ATCGTGAATT CAGGTCTGAT AAGCCCTGAA GGACTTGCCA TAGACACAT CCGCAGAAC 3600
ATGTACTGGA CCGACAGTGT CCTGGATAAG ATAGAGAGCG CCCTGCTGGA TGGCTCTGAG 3660
CGCAAGGTCC TCTTCTACAC AGATCTGGTG AATCCCCGTG CCATGCTGT GGAATCAATC 3720
CGAGGCAACT TGTACTGGAC AGACTGGAAT AGAGAAGCTC CTAAGATTGA AACGTCTCT 3780
TAGAATGAG AAGACAGAA AATTCTGATC AATACAGACA TTGATTGACC CAATGGCTTA 3840
ACCTTGAACC CTTTCTCTAA ACTGCTCTGC TGGGCAGATG CAGGAACCAA AAACTGGAG 3900
TGTAACCTAC CTGATGGAAC TGGACGGCGT GTCATTCAA ACAACCTCAA GTACCCCTTC 3960
AGCATCGTAA GCTATGCAGA TCACCTCTAC CACACAGACT GGAGGAGGGA TGGTGTGTA 4020
TCAGTAATA AACATAGTGG CCACTTTACT GATGAGTATC TCCAGAACCA ACGATCTCAC 4080
CTCTACGGGA TAATGCACT CTACCCCTAC TGCCCAACAG GAAGAAAGTA AGTACAGTAA 4140
TGTAAGGAA GACTTGGAGT TTACAATCAG AACCTGGACC CTAAGAACCA GTGACTGCAA 4200
AGGCAAGAA AGTAAAAAAG GAATTGGCCA TTAGACGTT CTGAGCATCC AAGATGAACA 4260
TTTTGTAGT CAAAAAGACT TTTGTAAAA GCTGATACCT CAATCTTTAC TACTGTATTT 4320
TTAAAAATGA AGGTTGTTAT TGCAAGTTTA AAAAGGTAAC AGAATTTTA CTGTGCTTA 4380
TTAAGCAAC TCTTGTGAAA CATTTATCAT TAATATTTAA AAGATCAAT TCATTCAACT 4440
YNAANKTCB GTTTAAGACT CTAAACCTGA TTTTGGCAT GGATTCCTTC TGGCCAAGAA 4500
ATTAAAGCAC ATGTGATCAA TATAACAATA TAATCCTAAA CCTTGACAGT TGGAGAAGCC 4560
ATTCAGAAC TGTATGGAAA GACCAATTA TTTATAGTTT CCCAACAAA GTTCTAAGAT 4620
TTTTTACTCT TGATCAGTG CATTTCTATT TATATCAAAA GGTGCTAAAA TGATTCATTT 4680
HGSATTTTCT GATCTGTAG TGCTCTATA GAAGTACCCA CAGAAAGTAA AGTATCACAT 4740
TTATAAATAC CAAAGATGTA ACAATTTTAA AATTTCTAG ATTACTCCAA TAAAGTGTTT 4800
TAAGTTTAAA AAAAAAAAAA AAAAAAAAAA

Seq ID No: 117 Protein sequence
Protein Accession #: NP_031387.1

50
55
60
65
70
75

1 11 21 31 41 51
MEGDRVAGRP VLSSLFVLLL LQLMLRAAA LHPDELFPHG ESWDQLLQE GDDVKLSRGE 60
AGESPALLTK PDSATSTWAP TASSPLRTSP GKRSMTMIS PPTSRLPPLF WRTSTRATAE 120
AESCTERTFP PQCWAMPFM CALASRALRA FYFHPRLPGH LGARRLRGG QTRALPSGEL 180
NTFQAVLADG GSDSYALFLY PANGLQFLGT RPKESSYNVQL QLPARVGFPC GEADDLKSEG 240
FYPSITSTEQ SVIKLYQLSN LGIPGVWAFH IGSTSPLDNH RPAAVGDLA AHSSVPLGRS 300
PSHATALESD YNEDNLDYD VNEBEAEVLP GEPEBALNGH SSIDVSPQSK VDTKPLESS 360
TLDPHTKEGT SLGEVGGPDL KQQVEPWDER ETRSPAPPEV DRDSLAPSWB TPPFPYENG 420
IQPYPDGQPV PSEMVPFPAH PEEHVLRSY PASGHTTPLS RGTVEVGLED NIGSNTVEFT 480
YNAANKTCB HNRQCSRHA PCTDYATGPC CHCQSKFYGN GKHLPEGAP HRVNGKVSGH 540
LHVGHTPVHF TDVDLHAYIV GNDGRAYTAI SHIPQPAQA LLPLTPIGGL FGWLFPALEKP 600
GSENGFSLAG AAPTHMEVT FYPGETVRI TQTAEGLDPE NYLSIKTNIQ GQVPPVPANF 660
TAHISPYKEL YHYSdstVTS TSSRYSLTF GAINQWTSYR IHQNIYQVC RHAPRHPSFP 720
TTQQLNVDRV FALYNDEERV LRFAVTNQIG PVKEDSDPTP VNPCYDGSIM CDTTARCHPG 780
TGVDYTCBCE SGYQGDGRNC VDENECATGF HRCGFNSVCI NLPGSYRCEC RSGYEFADDR 840
HTCILITPPA NPCEGDGSHC APAGQARCVH HGGSTFSCAC LPGAAGDGHQ CTDVDECSGN 900
RCHPAATCYN TPGFSRCRCQ PGYVGDGFCQ IPDSTSSLTP CBQQQRHAQA QYAYPGARFH 960
IPQCDGQGNF LPLQCHGSTG FCWCDVDPGH EVPGTQTTPG STFPFCGSP EPTQRFPPTIC 1020
ERWRENLEH YGGTTRDDQY VPQDDDLGHF IPLQCHGKSD FCWCDKDRG EVQGTRESQPG 1080
TTPACIPTVA PPMVRPTPRP DVTTPSVGTF LLYTQGGQIG YLPLNGTRLQ KDAAKTLLESL 1140
HGSIIIVGIDY DCRERMVYWT DVAGRTISRA GLELGAEPET IVNSGLISPE GLAIDHIRRT 1200
MYWTDVLDK IESALLDGSE RKVLFYTDLV NPRAIAVDPI RGNLYWTDWN REAPKLETSS 1260
LDGENRRILI NTDIGLPNGL TFDPFKLLC WADAGTKKLE CTLPDGTGRR VIQNNLKYPF 1320
SVSYADHPY HTDWRDGVV SVNKHSQOFT DEYLPQRSH LYGITAVYPT CPTGRK

Seq ID No: 118 DNA sequence
Nucleic Acid Accession #: NM_003088.1
Coding sequence: 112-1593 (underlined sequences correspond to start and stop codons)


```

1      11      21      31      41      51
|      |      |      |      |      |
5  GCGAGGGTG CGTGGGGCC GCGGCAGCCG AACAAAGGAG CAGGGGCGCC GCCGCAGGGA 60
   CCGCCACACC ACCTCCGGG GCGGCGCAGC GGCTCTCGT CTACTGCCAC CATGACCGCC 120
   AACGGCACAG CGAGGGCGGT GCAGATCCAG TTGGGCTCA TCAACTCGG CAACAAGTAC 180
   CTGACGGCGG AGCGTTTCGG GTTCAAGGTG AACGGTCCG CCAGCAGCCT GAAGAAGAAG 240
   CAGATCTGGA CGCTGGAGCA GCCCCTGAC GAGGCGGGCA GCGCGGCGGT GTGCTGCGC 300
   AGCCACCTGG GCGCTACCT GCGGCGGAC AAGGACGGCA ACGTGAACCT GAGCGCGAG 360
   GTGCCCGTTC CCGACTGCCG TTCTCTCATC GTGGGCGACG ACGACGCTG CTGCTCGCTG 420
   CAGTCCGAGG CGCACGGGCG CTACTTCGGC GGCACGAGG ACCGCTGTG CTGCTTCGG 480
   CAGACGGTGT CCGCGCGCGA GAGTGGAGC GTGCACATCG CCATGCACCC TCAGGTCAAC 540
   ATCTACAGTG TCACCCGTAA GCGCTACGCG CACCTGAGCG CGCGCGCGCG CGACGAGATC 600
   GCGCTAGAGC GCGACGTGCC CTGGGGCGTC GACTCGCTCA TCACCTTCGC CTTCACGAGC 660
   CAGCGCTACA GCGTGCAGAC GCGCGACAC CGCTTCTGCG GCCACGACGG GCGCTCGTGT 720
   CAGCGCGCCG AGCGGGCCAC TGGCTACACG CTGAGTTCCG GCTCGGCAA GGTGCGCTTC 780
   CGCGACTGCG AGGGCGGTGA CCTGGCGCGG TCGGGGCCCA GCGGCACGCT CAAGGGCGGG 840
   AAGCCACAGC AGGTGGGCAA GAGCAGGCTC TTGTCTTGG AGCAGAGCTG CGCCCAAGTC 900
   GTGCTCAGG CGGCAACGGA GAGGAACGTG TCCACGCGCC AGGGTATGGA CTGTCTGCC 960
   AATCAGTACG AGGACACGGA CCAGGAGACC TTCCAGCTGG AGATCGACCG CGACCCAA 1020
   AAGTGTGCTC TCGTACCCA CACGGGCAAG TACTGGACCG TGACGGCCAC CGGGGGCGTG 1080
   CAGTCCACCG CCGTCCAGCA GAATGCCAGC TGCTACTTTG ACATCGAGTG GGTGACCGG 1140
   GCGATCACAC TGAGGGCGTC CAATGGCAAG TTGTGACCT CCAAGAAGAA TGGGCACTG 1200
   GCGCGCTCGG TGGAGACAGC AGGGGACTCA GAGCTCTTCC TCATGAAGCT CATCAACCG 1260
   CCGATCATCG TGTTCGCGG GGAGCATGGC TTCTCGGCT CCGCAAGGT CAGGGGCAAC 1320
   CTGAGCGCCA ACGCTCCAG CTATGACGTC TTCCAGCTGG AGTTCAACGA TGGCGCTTAC 1380
   AACATCAAAG ACTCCACAGG CAAATACTGG ACGTGGGCA GTGACTCCG GGTACCCAG 1440
   AGCGGGGACA CTCCTGTGGA CTCTCTTTC GAGTCTGCG ACTATAACAA GGTGCCATC 1500
   AAGTGGGCGG GCGCTACCT GAAGGGGAC CACGAGGCG TCCTGAAGGC CTGGCGGAA 1560
   ACCGTGAGCC CGGCTCGCT CTGGGAGTAC TGGGGCGGCG CGCTCTTCC CGGCGCTGC 1620
   CCACATGGCG GCTCTGCGA ACCCTCGCTG CTAACCCCTT CTCGCGCAGG TGGGCTCCAG 1680
   GCGGGGAGCG AAGCCCCCTT GCCTTTCAAA CTGGAACCC CAGAGAAAC GGTGCCCCA 1740
   CCTGTGCGCC CTATGAGCTC CCACTCTCC CCGCGCGCG GTTCCCTAC TCCCTCGGG 1800
   TCAGCGGCTG CGGCTGCGC CTGGGAGGGA TTTCAGATGC CCGTCCCTC TTGTCTGCA 1860
   CGGGGCGAGT CTGGCACCTC TTTCTCTGA CCTCAGACCG CTCTGAGCCT TATTCTCTG 1920
   GAAGCGGCTA AGGACCGTGT GGGGCGTGG AGCCCTGGGC GTGTAGTGT ACTGGAATCT 1980
   TTGTCTCTC CCAGCCACCT CCTCCAGCC CCGCAGGAGA GCTGGGACA TGTCCCAAG 2040
   CTGTGAGTGG CCTCCCTGG TGCATGTCC CGAAACCCC TGCTTGGGA GGAAGCTGT 2100
   CGGAGGGCTG AGGACTGACC CTGTGTGTGT TTTTGTGGT GGTGGCTGGA AACAGCCCT 2160
   CTCCACGCTG GGAGAGGCTC AGCCTGGCTC CCTTCCCTGG AGCGGACGG GGTGACGGC 2220
   ACAGGGTCTG CCGCTGCAC GTTCTGCCAA GGTGGTGGTG GCGGGCGGGT AGGGGTGTGG 2280
   GGGCGTCTT CTCTCTGTCT CTTCCTTTC ACCCTAGCCT GACTGGAAG AGAAAATGAC 2340
   CAAATCAGTA TTTTCTTAA TGAATATTA TTGCTGAGG CGTCCAGGC AAGCCTGGCT 2400
   GTAGTAGCGA GTGATCTGGC GGGGCGGCTC TCAGCACCT CCGCAGGGG TGCACTTCAG 2460
   CCGCTCTTCT CGTCTCTTCC CGTCCAGCCC CAGCCTGGG CCGGGCTGC GACACCTGG 2520
   GCCAGAGCCC CTGCTGTGAT TGGTGTCTCC TGGGCTCCC GGGTGGATGA AGCCAGCGT 2580
   GCGCCCTCC GGGAGCCCTG GGGTGGAGCG CCGGGCCCCC CCGCTGCCA GCCTCCCCG 2640
   TCCCCAACAT GCATCTCACT CTGGGTGTCT TGGTCTTTA TTTTGTGAA GTGTCAATT 2700
   TATAACTCTA AACGCCATG ATAGTAGCTT CAAACTGGAA ATAGCGAAT AAAATAACTC 2760
   AGTCTGC

```

Seq ID No: 119 Protein sequence:

Protein Accession #: NP_003079.1

```

55 1      11      21      31      41      51
   |      |      |      |      |      |
   MTANGTAEAV QIQFGLINCQ NKYLTAFAFG FKNVASASSL KKKQIWTLEQ PPDAGSAAV 60
   CLRSHLGRYL AADKDGNVTC EREVPGPDCR FLIVAHDDGR WSLQSEAHRR YFGTEDRLS 120
   CFAQTVSPA E KNSVHIAMHP QVNIYSVTRK RYAHLSARPA DEIAVD RVP WGVDSLITLA 180
   PQDQRYSVQT ADHRFLRHG RLVARPEPAT GTTFEPRSGK VAFRDCRGY LAPSGSGTL 240
   KAGKATKVGK DELFALEQSC AQVVLQANE RNVSTRQGM LSANQDETD QETFLQLIDR 300
   DTKKACPRTH TGKYWTLTAT GGVVQSTASS NASCYFDIEW RDRRITLRAS NGKPVTSKKN 360
   GQLAASVSTA GDSEFLMKL INRPILVFRG EHGPIGCRKV TGTLDANRSS YDVFQLEFND 420
   GAYNIDSTG KYWTVGSDSA VTSSGDTVPD FFFBPCDYNK VAIKVGGRYL KGDHAGVLKA 480
   SAETVDPASL WEY

```

Seq ID NO: 120 DNA sequence

Nucleic Acid Accession #: NM_006404.1

Coding sequence: 25-741 (underlined sequences correspond to start and stop codons)

```

75 1      11      21      31      41      51
   |      |      |      |      |      |
   CAGGTCCGGA GCCTCAACTT CAGGATGTTG ACAACATTGC TGCCGATACT GCTGCTGTCT 60
   GGCTGGGCTT TTGTAGCCA AGACGCTCA GATGGCTCC AAAGACTTCA TATGCTCCAG 120
   ATCTCTACT TCCGCGACCC CTATCACGTG TGTACACAG GCAACGGCTC GCTGGGGGA 180

```

5 CACCTAACGC AGGTGCTGGA AGGCCCCAGAC ACCAACACCA CGATCATTTCA GCTGCAGCCC 240
 TTCCAGGAGC CGAGAGGCTG GCGCGGCACG CAGAGTGGCT TGCACTCCTA CCTGCTCCAG 300
 TTCCAGGAGC TGGTGGGCTG GGTGACACAG GAGGGGACCT TGGGCTTTCC TCTGACCATC 360
 CGCTGCTTCC TGGGCTGTGA GCTGCCCTCC GAGGGCTCTA GAGCCCATGT CTCTTGGAA 420
 GTGGCTGTGA ATGGGAGCTC CTCTTGTGAGT TTCCGGCCGG AGAGAGCCTT GTGGCAGGCA 480
 GACACCCAGG TCACCTCCGG AGTGGTCACC TTCACCTGTC AGCAGCTCAA TGCCTACAAC 540
 CGCACTCGGT ATGAACCTGG GGAATTCCTG GAGGACACCT GTGTGCAGTA TGTGCAGAAA 600
 CATATTTCCG CGGAAACAC GAAAGGGAGC CAAACAAGCC GCTCTACAC TTCGCTGTC 660
 CTGGGCGTCC TGGTGGGCGG TTTCATCATT GCTGGTGTGG CTGTAGGCAT CTCTCTGTGC 720
 10 ACAGTGGGAC GGCAGATGTA ATTACTCTCC AGCCCGGTCA GAAGGGGCTG GATTGATGGA 780
 GGCTGGCAAG GGAAGATTTT AGCTCAGTGT GAAGCCAGAC TCCCAACTGT AAACACCAGA 840
 AGGTTTGGAG TGACAGCTCC TTTCTTCTCC CACATCTGCC CACTGAAGAT TTGAGGGAGG 900
 GAGATGGGAG AGGAGAGGTG GACAAAGTAC TTGGTTTGTCT AAGAACCCTAA GAACTGTAT 960
 GCTTTGCTGA ATTAGTCTGA TAAGTGAATG TTTATCTATC TTTGTGGAAA ACAGATAATG 1020
 15 CAGTGGGCGC AGGAGCCCTA TGGGCCATCC TCCAAAGACA GACAGAAATCA CCTGAGGCGT 1080
 TCAAAAGATA TAAACAATAA AACAAATCAT CCACAATCAA AATACAACAT TCAATATCTC 1140
 CAGGTGTGTG AGACTTGGGA TGGGACGCTG ATATAATAGG GTAGAAAGAA GTAACACGAA 1200
 GAGTGGTGG AAAATGTAAA TCCAAGTCAT ATGGCAGTGA TCAATTATTA ATCAATTAAT 1260
 20 AATATTATAA AATTTCTTAT ATTT

Seq ID No: 121 Protein sequence
 Protein Accession #: NP_006395.1

25 1 11 21 31 41 51
 MLTTLPIILL LSGWAFCSQD ASDGLQLRHM LQISYFRDPY HWVYQGNASL GGHLTHVLEG 60
 PDITMTIQL QPLQSPESWA RTQSLQLSYL LQPHGLVRLV HQERTLAPFL TIRCFGLCEL 120
 PPEGSRHHVF FEVAVNGSSF VSPFRERALW QADTQVTSQV VTFTLQQLAA YNRTRYELRE 180
 30 FLEDTQCVQV QKHISAENTK GSQTSRSYTS LVLGLVLGGF IIAGVAVGIF LCTGRRRC

Seq ID NO: 122 DNA sequence
 Nucleic Acid Accession #: none found
 Coding sequence: 2-505 (underlined sequences correspond to start and stop codons)

35 1 11 21 31 41 51
 CGAGAAGCTG GGAGAGACAC CACTTGTCCC TGAACAAGAC AATTGAGTAA CATCTATTCC 60
 TGAGATTTCCT CGATGGGGAT CACAGAGCAC GATGCTTACC CTTCAAATGT CCTTCAAGC 120
 40 CGAGTCAAGG GCCACTATCA CCCCATCAGG GAGCGTGATT TCCAAATTTA ATTCTACGAC 180
 TTCTCTGCGC ACTCTCAGG CTTTGGACTC CTCTCTGCCC GTGGTCTTCA TATTTGTGAG 240
 CACAGCAGTA GTAGTGTGG TGATCTTGAC CATGACAGTA CTGGGGCTTG TCAAGCTCTG 300
 CTTTACGAA AGCCCTCTT CCCAGCCAAG GAAGGATCT ATGGGCGCGC CGGGCCTGGA 360
 45 GAGTGATCTG GAGCCCGCTG CTTTGGGCTC CAGTTCTGCA CATTCACAAA ACAATGGGGT 420
 GAAAGTCGGG GAGTGTGATC TGGGGGACAG AGCAGAGGGT GCCTTGCTGG CGGAGTCCCC 480
 TCTTGGCTCT AGTGATGCAT AGGGAACAG GGGACATGGG CACTCTGTGG AACAGTTTCT 540
 CACTTTGATG GAAACGGGGA ACCAAGAGGA ACTTACTTGT GTAAGTACA ATTCTGCGAG 600
 AAATCCCCCT TCCTTAATAT TCCTTTTACT CCCTGAGGA GCTAAATCAG AACTGCACAC 660
 50 TCCTTCCTCG ATGATAGAGG AAGTGAAGT GCCTTTAGGA TGGTGATCT GGGGGACGG 720
 GTAGTGTGGG GGAGAGATAT TTTCTTATGT TTATTGGAG AATTGGAGA AGTGATTGAA 780
 CTTTTCAGAA CATTTGAAAC AAATAGAAC CAATATAATT TACATTAAAA AATAATTCT 840
 ACCAAAATGG AAAGGAAATG TTCTATGTTG TTCAGGCTAG GAGTATATTG GTTCGAAATC 900
 CCAAGGAAAA AAATAAAAT AAAAATTAA AGGATTGTTG ATAAAA

55 Seq ID No: 123 Protein sequence
 Protein Accession #: none found

60 1 11 21 31 41 51
 EKLGETPLVP EQDNVSITSIP EIPRWGSQST MSTLQMSLQA ESKATITPSG SVISKFNSTT 60
 SSATPQAFDS SSAVVFIFVS TAVVVLILT MTVLGLVKLC FHESPSSQPR KESMPPGLE 120
 SDPBPAAAGS SSAHCTNGV KVGDCDLRDR AEGALLAESF LGSSDA

65 Seq ID NO: 124 DNA sequence
 Nucleic Acid Accession #: NM_006500.1
 Coding sequence: 27-1967 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 ACTTCGCTCT CGCCCTCCGG CCAAGCATGG GGCCTCCAG GCTGGTCTGC GCCTTCTTGC 60
 TGGCCGCTCT CTGCTGCTGT CCTCGCTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
 CGCCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCTCTCC 180
 75 AGTCCAGG CAACCTCAGC CATGTGCACT GTTTTCTGT CCACAAGGAG AAGCGGACGC 240
 TCATCTTCCG TGTGCGCAG GCCAGGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGGC 300
 TCAGCTCCA GGACAGAGG GCTACTCTGG CCTGACTCA AGTACCCCC CAAGACGAGC 360

	GCATCTTCTT	GTGCCAGGGC	AAGCGCCCTC	GGTCCCAGGA	GTACCGCATC	CAGCTCCGGG	420
	TCTACAAGC	TCCGGAGGAG	CCAAACATCC	AGGTCAAACC	CCTGGGCATC	CCTGTGAACA	480
	GTAAGGAGCC	TGAGGAGGTC	GCTACCTGTG	TAGGGAGGAA	CGGGTACCCC	ATTCTCTAAG	540
5	TCATCTGGTA	CAAGAAAGGC	CGGCTCTGTA	AGGAGGAGAA	GAACCGGGTC	CACATTCAAT	600
	CGTCCCAGAC	TGTGGAGTCG	AGTGGTTTGT	ACACCTTGCA	GAGTATTCTG	AAGGCACAGC	660
	TGGTTAAAGA	AGACAAAGAT	GCCCAAGTTT	ACTGTGAGCT	CAACTACCGG	CTGCCCAATG	720
	GGAAACCAT	GAAGGAGTCC	AGGGAAGTCA	CGTCCCCTGT	TTTCTACCGG	ACAGAAAAGG	780
	TGTGGCTGGA	AGTGGAGCCC	GTGGGAATGC	TGAAGGAAGG	GGACCGGCTG	GAATTCAGGT	840
10	GTTTGGCTGA	TGGCAACCTT	CCACCACTCT	TCAGCATCAG	CAGCAGAAC	CCCAGCACCA	900
	GGGAGGCAGA	GGAAGAGACA	ACCAACGACA	ACGGGGTCTT	GOTGCTGGAG	CCTGCCCGGA	960
	AGGAACACAG	TGGGCGCTAT	GAATGTTCAGG	CCTGGAACTT	GGACACCATG	ATATCGCTGC	1020
	TGAGTGAACC	ACAGGAACCTA	CTGTGAACT	ATGTGTCTGA	CGTCCGAGTG	AGTCCCGCAG	1080
	CCCTGTGAGG	ACAGGAAGGC	AGCAGCCTCA	CCCTGACCTG	TGAGGAGAGG	AGTAGCCAGG	1140
	ACCTCGAGTT	CCAGTGGCTG	AGAGAAGAGA	CAGACCAAGT	GCTGGAAGAG	GGGCTGTGTC	1200
15	TTCAAGTTGCA	TGACCTGAAA	CGGGAAGCAG	GAGGCGGCTA	TCGCTGCTG	CGCTCTGTGC	1260
	CCAGCATACC	CGGCTGAAC	CGCACACAGC	TGGTCAAGCT	GGCCATTTT	GGCCCCCTT	1320
	GGATGGCATT	CAAGGAGAGG	AAGGTGTGGG	TGAAGAGAA	TATGTTGTTG	AATCTGTCTT	1380
	GTGAAGCGTC	AGGGCACCCC	CGGCCACCA	TCTCTGGA	CGTCAACGGC	ACGGCAAGTG	1440
	AACAGACCA	AGATCCACAG	CGAGTCTCTA	GCACCTGAA	TGTCTCTG	ACCCGAGGAG	1500
20	TGTTGGAGAG	AGGTGTTGAA	TGCAAGCGCT	CCACGACCT	GGGCAAAAC	ACCAGCATCC	1560
	TCTTCTGGA	GCTGCTCAAT	TTAACACCC	TCACACAGA	TCCAAACACA	ACCAGTGGCC	1620
	TCAGCACTTC	CACTGCCAGT	CCTCATACCA	GAGCCAAACG	CACCTCCACA	GAGAGAAAGC	1680
	TGCCGAGGCC	GGAGAGCCGG	GGCGTGTCA	TGCTGGCTGT	GATTTGTGTC	ATCCTGTGTC	1740
25	TGGCGGTGCT	GGCGCTGTCT	CTCTATTTC	TCTATAAGAA	GGGCAAGCTG	CGGTGCAGGC	1800
	GCTCAGGAA	GCAGGAGATC	ACGCTGCCCC	CGTCTGTGAA	GACCGAACTT	GTAGTTGAAG	1860
	TTAAGTCAGA	TAAGCTCCCA	GAAGAGATGG	CGCTCTGCA	GGGCAGCAGC	GGTGACAAGA	1920
	GGGCTCCGGC	AGACAGGGGA	GAGAAATACA	TGATCTGAG	GCAITAGCCC	OGAATCACTT	1980
	CAGCTCCCTT	CGCTGCTGG	ACCAITCCCA	CGTCCCTGCT	CACCTTCTCT	TCAGCCAAAG	2040
	CCTCCAAAGG	GACTAGAGAG	AAGCCTCCCT	CTCCCTTCAC	CTGCACACCC	CCTTTCAGAG	2100
30	CGCCAGTCAGG	TTAGGACCTG	AGGACCTCAC	TGGCCCTGCG	AAGCGGCTTT	TCAGGGACCA	2160
	GTCCACCAACC	ATCTCTCTCA	CGTTGAGTGA	AGCTCATCCC	AAGCAAGGAG	CCCCAGTCTC	2220
	CCGAGCCGGT	AGAGAGGTTT	CTTCCAGAAC	GTGTTTTTTC	TTTACACACA	TTATGGCTGT	2280
	AAATACCTGG	CTCTGCCAG	CAGCTGAGCT	GGGTAGCCTC	TCTGAGCTGG	TTTCTGCCCC	2340
35	CAAGGGCTGG	CTTCCACCAT	CCAGGTGCAC	CACCTGAAGT	AGGACACACC	GGAGCCAGGC	2400
	CGCTGCTCAT	GTTGAAGTGC	GCTGTTTACA	CCCGCTCCGG	AGAGCACCCC	AGCGGCATCC	2460
	AGAAGCAGCT	GCAGTGTGTC	TGCCACCAAC	CTCTGCTCG	CCTCTTCAA	GTCTCTCTGT	2520
	ACATTTTTTC	TTTGGTCAGA	AGCCAGGAAC	TGGTGTCAIT	CCTTAAAGA	TACGTCCGG	2580
	GGCCAGGTGT	GGTGCTCAC	GCCTGTAAAT	CCAGCACTTT	GGGAGGCCGA	GGCGGGCGGA	2640
40	TCACAAAGTC	AGACAGAGAC	CATCTGGCT	AACACGGTGA	AACCTGTCT	CTACTAAAAA	2700
	TACAAAAAAA	AATTAGCTAG	GCGTAGTGGT	TGGCACTAT	AGTCCAGCT	ACTCGGAAGG	2760
	CGTCAAGCAG	AGAAGGTAT	GAATCCAGGA	GGTGGAGCTT	GCAGTGAGCC	GAGACCGTGC	2820
	CACCTGCACT	CAGCTCGGCC	AACACAGCGA	GACTCCGTCT	CGAGGAAAAA	AAAAGAAAAA	2880
	ACGCGTACCT	CGCGTGAGGA	AGCTGGGCGC	TGTTTTGAG	TTCAAGTGAA	TTAGCCTCAA	2940
45	TCCCGGTGTT	CACCTGTCTC	CATAGCCCTC	TGATGGATC	ACGTAAAACT	GAAGGCAGC	3000
	GGGAGGACAG	CAAGATGAG	GTCTACACTG	TCCTTATGCG	GGATTAAAGC	TATGGTTATA	3060
	AGTCAACCAA	ACTTCTACAA	ACCAAGCTCA	GGGCCCCAAC	CCTAGAAGGG	CCCAATGAG	3120
	AGAATGTGAT	TTAGGATGAG	AAAAAGGGGC	CTGCGTAGAG	CTTCGGGTGT	GTGTGTCTGT	3180
	CTGTGTGTAT	GCATACATAT	GTGTGTATAT	ATGGTTTTGT	CAGGTGTGTA	AATTTGCAAA	3240
50	TGTGTTCTCT	TATATATGTA	TGTATATATA	TATATGAAAA	TATATATATA	TATGAAAAAT	3300
	AAAGCTTAAT	TGTCCAGAA	AATCATACAT	TGCTTTTTTA	TTCTACATGG	GTACCAACAG	3360
	AACCTGGGGG	CCTGTGAAAC	TACAACCAAA	AGGCACACAA	AACCGTTTCC	AGTTGGCAGC	3420
	AGAGATCAGG	GGTACCTCT	GCTTCTGAGC	AAATGGCTCA	AGCTCTACCA	GAGCAGACAG	3480
	CTACCTACTT	TTTCAGCAGC	AAAACGTCCC	GTATGACGCA	GCACGAAGGG	CCTGGCAGGC	3540
55	TGTTAGCAGG	AGCTATGTCC	CTTCTATCG	TTTCCGTCCA	CTT		

Seq ID No: 125 Protein sequence
 Protein Accession #: NP_006491.1

60	1	11	21	31	41	51	
	MGLPRLVCAF	LLAACCCCP	VAGVPGEAEQ	PAPELVEVEV	GSTALLKCGL	SQSQGNLSHV	60
	DWFSVHKKEK	TLFVRVRQGG	GOSEPGYEYEQ	RLSLQDRGAT	LALTQVTFQD	ERIFLCQGR	120
	PRSQEYRIQL	RVYKAPBEPH	IQVNPLGIPV	NSKEPEEVAT	CVGRNVPIPI	QVIWYKNGRP	180
65	LKEEKNRVHI	QSSQTVESGG	LYTLQSLILKA	QLVKEDKDAQ	FYCHLWYRLP	SGNHMKESRE	240
	VTVPVFPTE	KVWLEVEFVG	NLKEGDRVEI	RCLADGNPPP	HFSISKQNP	TREAEETTN	300
	DNGVLVLEPA	RKEHSGRYEC	QAWNLDTMIS	LLSBPQBLV	NYVSDVRVSP	AAPBRQEGSS	360
	LTLTCEAES	QDLFPQLRE	ETDQVLERGP	VLQHLDLKRE	AGGGYRCVAS	VPSIPGLNRT	420
70	QLVKLAIFGP	PWMAFKERKV	WVKENMVLNL	SCBASGHPRP	TISWNVNGTA	SEQDQDPQRV	480
	LSTLNLVITP	ELLETVGEVCT	ASNDLSKNTS	ILFLELVNLT	TLTPDSNTTT	GLSTSTASPH	540
	TRANSTSTER	KLPEPESRGV	VIVAVIVCIL	VLAVLGAIVL	FLYKKGKLP	RRSGKQBITL	600
	PSRKTELTV	EVKSKLPBZ	MGLLQSSGD	KRAPDQGEK	YIDLKH		

75 Seq ID NO: 126 DNA sequence
 Nucleic Acid Accession #: NM_001955.1

	1	11	21	31	41	51	
5	GGAGCTGTTT	ACCCCCACTC	TAA TAGGGGT	TCAATATAAA	AAGCGCGCAG	AGAGCTGTGCC	60
	AAGTCAGACG	CGCCTCTGCA	TCTGCGCGCAG	GAGCAGCGGT	TCTGCGCCTC	CTCGCACTGCC	120
	AGGCTCTCAC	CACCGCGCGC	TGCGCTGTGA	CGCGCTCGCG	TCGTCTCTGC	CTCTCTGGGC	180
	AGGCGCTGCC	TTTTCTCCCC	GT TAAAGGGC	ACTTGTGGCTG	ATGATGCTCT	TTGAGATCTG	240
10	AGGAAAGCGC	AGCGCTTGTA	GGGACACTGA	GCGTGTTTTC	TTGCTTTTCC	TTTGCGTTCA	300
	GT TTGACGCG	GAGGTTTTTTG	ATCCCTTTTT	TTCAGAAATG	ATTATTGTTC	CATGATTTTC	360
	TC TCTGCTGT	TTTGTCGCTT	CAAGAGGACT	CCAGAAACAG	CAGTCTTAGG	CGCTGAGCTC	420
	ACGCGCGTGG	GTGAGAAAGC	CGGGGAGAAA	CCCACTGCCA	GTCACCGCTG	CGGCTCGCCG	480
	GGCTCCAAAG	CGTCTCTCTG	CTCGTCCCTG	CGATGTAAG	ATGCTGTCTA	CTGCTTCGAC	540
15	CTGGAACATCA	TTTGGGTGCA	CAC TCCCGAG	CAGCTGTGTC	CGATATGACT	TGGAAGACCT	600
	ATGCACAGA	GAGCCTTGGA	GAA TTTACTT	CCCAACAAGG	CAATACACCG	TGGAATATCA	660
	TGCCATATGT	CTAGCCAAAA	AGCAAGAAGT	TGCTGGAATT	TTTCCCAAGC	AGGAAAGAAA	720
	CTCAGGGCTG	AAGCACTATG	GAGAAAGATG	TGGAAATATC	ATAAGAAAGG	AAAGACATGT	780
20	TCCAAAGCTG	GGAAAAAGTG	TATTTATPAC	CAGTATATGTA	GAGGAGAGAA	AATTCAGAAAG	840
	AGTTCAGAGG	AACACCTAAG	ACAAACACAG	TCGGAGACCA	TGGAAGAACAC	CGTCAAAATCA	900
	TCTTTTCACT	ATCCCAAGCT	GAAAGGCGAG	CTCCACCAAG	AGCGTTATGT	GACCCCAACAC	960
	CGAGCACATT	GGTGAACAGC	TTCGGGGCCT	GCTCTGAAGC	ATAGCTTCCA	CGGAGAGCCC	1020
	TGTGGCGCAG	TCTGCACTCT	CCACCCGTGC	TGGGATCAGA	CGAGGAGCAT	CTCTGTCTGG	1080
25	TTCTGTAGTG	GCAAGGAGCT	AGCGTCTCTG	TTCARAACAT	TCCAGAAAGT	GTTAAGGAAGT	1140
	TCCCCAACAC	ATCTTCACTG	GCTCTTCATCA	GTGGTACACT	CTTTGTCTCT	TTCTTTTCATC	1200
	TGGGGTAGGC	AATGACAGCT	TCAGCAGAAA	CACCAAGTCA	CTATCGAATT	C	

30

	1	11	21	31	41	51	
35	MDYLLMIFSL	LFVACQGAPE	TAVLGAELSA	VGENGGEKPT	PSPPWRLRRS	KRCSCSSLMD	60
	KCVYFCHLD	IIWVNTPEHV	VPYGLGSPRS	KRALENLLPT	KATDRENRCQ	CASQKDKKCW	120
	NPCQAGKELR	AEDIMEKDOWN	NHKKGKDCSK	LGKKCIYQQL	VRGKRIRRS	EEHLRQTRSS	180
	TRNSVKSSEF	HDPELKGKPS	RRRYVTHNRA	HW			

40 Seq ID NO: 128 DNA sequence
Nucleic Acid Accession #: NM_001721.1
Coding sequence: 34-2061 (underlined sequences correspond to start and stop codons)

45	1	11	21	31	41	51	
	GCAAGCAGCG	AACAAGCTGA	GACGGATGAT	AAATATGGATA	CAAAATCTAT	TCTAGAAGAA	60
	CTTCTCTCA	AAAGATCAC	GCAAAAGAA	AAATGTCAC	CAAAATATTA	CAAGAAGCG	120
50	CTTTTGTGTT	TGACCACAAAC	AAACTCTTCC	TACTATGAT	ATGACAAAAT	GAAAAGGGGC	180
	ACGAGAAAGT	GATCCATTAC	AAATTAAGAA	ATCAGATGTG	TGGAGAAAGT	AAATCTCCAG	240
	GAGCAGACGC	CTGTAGAGAG	ACAGTACCCA	TTTCAGATTG	CTTATAAAGA	TGGGCTTCTC	300
	TATGTCTATG	CATCAAAATG	AGAGAGCGCA	AGTCAGTGGT	TGAAAGCATT	GAAAAGAGAG	360
	ATAAGGGGTT	ACCCCCAACC	CGTGTGCAAG	TACCATATGT	GGTCTTCCTG	GCAAGGGGAG	420
55	TTCTCTGTGT	GCCAGCAGAG	CTGTAAAGCA	GCCCCAGGAT	GTACCCCTGT	CGGACCATAT	480
	GCTAATCTGC	ATACTCGATG	CAATGAGAG	AAACACAGAG	TTCCCACTTT	CCAGACAGAA	540
	GTGCTGAAGA	TACCTCGGGC	AGATTCCTGT	CTCAAAATG	ATGCACCATC	TTTCAAGTAC	600
	ACTCTAGCCG	AATATGACAA	CGAATCAAGA	AAAGAACTGT	CGCTCCAGCC	ACCATCTTCA	660
	AGTACACAGT	TAGCGCAATA	TGACAGCAAC	TCAAAGAAAA	TCTATGGGTC	CCAGGCCAAC	720
60	TTCAACATCG	AGTATAGATG	AAAGGAAGAC	TTCCCTGACT	GGTGGCAAGT	AGAAAGAACT	780
	AAAAGTAGCA	GCAGCAGTGA	AGATGTTGCA	AGCAGTAACC	AAAAGAAAG	AAATGTGAAT	840
	CACACCACTT	CAAGAAATTTT	ATGGGAATCT	CCTGAGTACA	GTTCATCTGA	AGAGAGGAAA	900
	AACCTGGAGT	ATTATAGAGT	GTTTTCTGTT	AACATCTCCA	GATCACAATG	TGAACGAGTTA	960
	CTCAGACAAA	AGGGAAAAAG	AGGAGCAATT	ATTGTTAGAA	ATCTGAGCCA	AGTGGGAATG	1020
65	TACACAGTGT	CGTTATTATG	TAAAGCTGTG	ATGTGATAAA	AAGGAACATG	CAAACTATTAC	1080
	CAGTGTGATA	CAATGTCTGA	GAAACAAATTA	TACCTGGCAG	AAAACACTACT	TTTGTAACTT	1140
	TTTCCAAAGC	TTATCTATTCA	TCATCAACAC	AAATTACAGC	GCATGATCAC	ACGGCTCCGC	1200
	CACCTCTGTG	CAACAAAGGC	CAACCAAGGTC	CCCGACTCTG	TGTCCCTGGG	AAATGGAAAT	1260
	TGGGAAGTGA	AAAGAGAAGA	GATTAATCTT	TTGAAGAGAC	TGGGAAGTGG	CCAGTTTGGG	1320
70	GTGGCTCAGC	TGGGCAAGTG	GAAGGGGCAG	TATGATGTTG	CTGTTAAGAT	GATCAAGGAG	1380
	GGCTCCATGT	CAGAAGATGA	ATTCTTTACG	GAGGGCCAGA	TCTATGATGA	ACTCAGGCAT	1440
	CCCAAGCTGG	TAAATTTCTA	TGGAGTGTGT	TCAAAGGAAT	ACCCATATAT	CATAGTAGAT	1500
	GAATATATRA	GCAATGGCTG	CTTGTCTGAAT	TACTGTGAGG	GTACCGGAAA	AGGACTTGAA	1560
	CCTTCCAGCC	TTTAGAAGAT	GTGTCTAGAT	CGCTGTGAAG	GCATGCGCTT	TTGTGAGAGT	1620
75	CACCAATTCA	TACACCGGGA	CTTGGTGCTC	CGTAATCCTC	TGGTGGAGCA	AGATCTCTGT	1680
	GTGAAAGATG	CTGACTTTGG	ATGACCAAGG	TGTTTCTTGT	ATGACCAGTA	TGTCACTTGA	1740

5
 10
 15
 20
 25
 30

```

GTCGGAACAA AGTTTCCAGT CAAGTGGTCA GCTCCAGAGG TGTTCATTAA CTTCAAATAC 1800
AGCAGCAAGT CAGACGTATG GGCATTTGGG ATCCTGATGT GGGAGTGTGT CAGCCTGGGG 1850
AAGCAGCCCT ATGACTTGTA TGACAACTCC CAGGTGGTTC TGAAGGTCTC CCAGGGCCAC 1920
AGGCTTTACC GGGCCACCTT GGCATCGGAC ACCATCTACC AGATCATGTA CAGCTGCTGG 1980
CAGCAGCTTC CAGAAAAGCG TCCACATTTT CAGCACTCCG TGTCTTCCAT TGAACCACTT 2040
CGGGAAGAAAG ACAAGCATTTG AGAAGAAAT TAGGAGTGTCT GATAAGAAAT AATATAGATG 2100
CTGGCCAGCA TTTTCATTCA TTTAAGGAA AGTAGGAAGG CATAAGTAAT TTTAGCTAGT 2160
TTTAAATAGT GTTCTCTGTA TTGTCTATTA TTTAGAAATG AACAGGCAG GAAACAAAAG 2220
ATTCCCTTGA AATTAGATC AAATTAGTAA TTTGTTTTA TGCTGCTCCT GATATAACAC 2280
TTCCAGCCT ATAGCAGAAG CACATTTTCA GACTGCAATA TAGAGACTGT GTTCATGTGT 2340
AAAGACTGAG CAGAACTGAA AAATTACTTA TTGGATATTC ATTCTTTTCT TTATATTGTC 2400
ATTGTACAA CAATTAAATA TACTACCAAG TACAGAAATG TGGAAAAA AAACCG
  
```

Seq ID No: 129 Protein sequence:
 Protein Accession #: NP_001712.1

1 11 21 31 41 51
 20
 25
 30

```

MOTKSILEEL LLKRSQQKKK MSPNNYKERL FVLTKINLSY YBYDKMKRGS RKSSIEIKKI 60
RCVKKYNLEE QTPVERQYPP QIVYKDGILLY VYASNEESRS QWLKALQKEI RGNPHLLVKY 120
HSGFFVDGKF LCCQSCSKAA PGCTLWBAYA NLHTAVNEEK HRVPTFPDRV LKIPRAVPVL 180
KMDAPSSSTT LAQYDNESKK NYGSQPPSSS TSLAQVDSNS KKIYGSQPNF NMQYIPREDP 240
PDWQVVKLLK SSSSSSDVAS SNQKERNVNH TTSKISWEPP SSSSSSEERN LDDYDWFAGN 300
ISRSQSEQLL RQKKGKGAFF VRNSSQVQMY TVSLFSAKVN DKKGTVKHYH VHTNAENKLY 360
LAENYCFDSI FKLHYHQM SAGMITRLRH PVSTKANKVP DSVSLNGNIW ELKREBITLL 420
KELSGQFQV VQLGKWKQY DVAVKMKIEG SMSSEDFPQE AQTMKLSHP KLVPFYGVCS 480
KRYPIYIVTE YISNGCLLNY LRSHGKGLFP SLLLEMCYDV CEGMAFLSH QPHRDLAAR 540
NCLVDRDLCV KVSDFPMTRY VLDQYVSSV GTFKPVKWSA PEVPHYFKYS SKSDVMAFGI 600
LMWEVSLGK QPYLDYDMSQ VVLKVSQHR LYPHILASDT IYQIMYSCHW ELPEKRPFTQ 660
QLLSIEPLR EKDKH
  
```

Seq ID NO: 130 DNA sequence
 Nucleic Acid Accession #: NM_012072.2
 Coding sequence: 149-2107 (underlined sequences correspond to start and stop codons)

40
 45
 50
 55
 60
 65
 70
 75

```

AAAGCCCTCA GCGTTTGTGT CTTCTCTGCG GCGGAGTGG CTGCAGCTCA CCCCTCAGCT 60
CCCCTTGGGG CCCAGCTGGG AGCCGAGATA GAAGCTCCTG TCGCCGCTGG GCTTCTCGCC 120
TCCCGCAGAG GGCACACAG AGACCGGAT GGCACCTCC ATGGGCTGCG TGTCTGTGCT 180
GCTGCTGCTC CTGACCCAGC CCGGGGCGGG GACGGAGCT GACACGAGAG CGGTGGTCTG 240
CGTAGGGACC GCGTCTACAC CGGCCCACTC GGCCAAGCTG AGCGTGCAG AGGCCAGAA 300
CCACTGCAAC CAGAACGGGG GCAACCTGGC CACTGTGAAG AGCAAGGAGG AGGCCAGCA 360
CGTCCAGCGA GTACTGGCCC AGCTCCTGAG CCGGGAGGCA GCCCTGACCG CGAGATGAG 420
CAGTGTCTGG ATTGGGCTCC AGCGAGAGAA GGGCAAGTGC CTGACCCCTA GTCTGCGCT 480
GAAGGGCTTC AGCTGGGTGG GCGGGGGGGA GGACACGCTT TACTCTAACT GGCACAAGGA 540
GCTCCGGAAC TGTGTCATCT CCAAGCGCTG TGTGTCTCTG CTGCTGGACC TGTCCAGGCC 600
GCTCCTTCCC AACCGCTCCA CCAAGTGTGC TGAGGGCCCC TGTGGGAGCC CAGGCTCCCC 660
CGGAAGTAAC ATTGAGGGCT TCGTGTGCAA GTTCAGCTTC AAAGGCATGT CGCGGCTCT 720
GGCCCTGGGG GGCACAGGTC AGGTGACCTA CACCACCCCC TTCCAGACCA CCAATTCTCT 780
CTTGGAGGCT GTGCCCTTTG CTTCTGCGGC CAATGTAGCC TGTGGGGAAG GTGACAAGGA 840
CGAGACTCAG AGTCATTATT TCCTGTGCAA GGAGAAGGCC CCGATGTGT TCGACTGGGG 900
CAGCTCGGGC CCGCTCTGTG TCAGCCCCAA GTATGGCTGC AACTTCAACA ATGGGGGCTG 960
CCACAGGAGC TGCTTTGAAG GGGGGGATGG CTCTCTCCTC TCGGCTGCC GACCAGGATT 1020
CCGGCTGCTG GATGACCTGG TGACCTGTGC CTCTCGAAAC CTTGACAGT CCAGCCCATG 1080
TCTGGGGGGG GCCACGTGGG TCCTGGGACC CCATGGGAAA AACTACAGT GCGGCTGCCC 1140
CCAAGGTATC CAGCTGGACT CGAGTCAGCT GGACTGTGTG GACGTGGATG AATGCCAGGA 1200
CTCCCCCTGT GCCCAGGAGT GTGTCAACAC CCGTGGGGGC TTCCGCTGGG AATGCTGGGT 1260
TGGCTATGAG CCGGGCGGTC CTGGAGAGGG GGCTGTGAG GATGTGATG AGTGTGCTCT 1320
GGGTGCTCG CCTTGCAGCC AGGGCTGCAC CAACACAGAT GGCTCATTTT ACTGCTCTGT 1380
TGAGGAGGGG TACGTCTTGG CCGGGGAGGA CGGACTCAG TGCCAGGAGG TGGATGATG 1440
TGTGGGCCCC GGGGGCCCCC TCTGCGACAG CTTGTGCTTC AACACACAAG GTGCTTTCCA 1500
CTGTGCTGTC CTGCCAGGCT GGGTGTCTGC CCCAAATGGG GTCTCTTGCA CCAATGGGGC 1560
TGTGTCTCTG GACCAACCAT CTGGGCCCCC GATGAGGAG GACAAAGGAG AGAAAGAAAG 1620
GAGCACCGTG CCGCGCGCTG CAACAGCCAG TCCACAAAGG GGCCCGAGG GCACCCCAA 1680
GGCTACACCC ACCACAAGTA GACCTTCGCT GTCATCTGAC GCGCCCATCA CATCTGCCCC 1740
ACTCAAGATG CTGGCCCCCA GTGGGTCTCT AGGCGTCTGG AGGGAGCCCA GCATCCATCA 1800
CGCCACAGCT GCTCTGGCC CCCAGGAGCC TGCAGGTGGG GACTCTCTCG TGGCCACACA 1860
AAACAACGAT GGCATGACG GGCAAAAGCT GCTTTTATTC TACATCTAG GCAOCGTGGT 1920
GGCCATCTTA CTCTGTCTGG CCTTGGCTCT GGGGCTACTG GTCTATCGCA AGCGGAGAGC 1980
GAAGAAGGAG GAGAAGAAGG AGAAGAAGCC CCAGAATGGG CGACAGAGAT ACTCCTGGGT 2040
TCCAGAGCGA GCTGAGAGCA GGGCCATGGA GAACCAATAC AGTCCGACAC CTGGGACAGA 2100
CTGCTGAAG TGAGGTGGCC CTAGAGACAC TAGAGTCACC AGCCACCATC CTCAGAGCTT 2160
  
```

	TGAATCCCC	ATTCCAAAGG	GGCACCCACA	TTTTTTTGAA	AGACTGGACT	GGAACTTAG	2220
	CAAAACAATTG	TAAGTCTCCT	CCTTAAAGGC	CCCTTGAAC	ATGCAGGTAT	TTTCTACGGG	2280
	TGTTTGAATG	TCCTGAAGTG	GAAGCTGTGT	GTGGGCTGC	CACGGTGGGG	ATTTCTGTAC	2340
5	TCTATAATGA	TTGTTACTCC	CCCTCCCTTT	TCAAAATCCA	ATGTGACCAA	TTCCGGATCA	2400
	GGGTGTGAGG	AGGCTGGGGC	TAAGGGGCTC	CCCTGAATAT	CTTCTCTGCT	CACCTCCACC	2460
	ATCTAAGAGG	AAAAGGTGAG	TTGCTCATGC	TGATTAGGAT	TGAAATGATT	TGTTTCTCTT	2520
	CCTAGGATGA	AACTAAATC	AATTAATAT	TCAATTAGGT	AAGAAGATCT	GGTTTTTTGG	2580
	TCAAGGGAA	CATGTTGCGA	CTGGAAACAT	TTCTTTACAT	TTGCATTCTC	CCATTTCGCC	2640
10	AGCACAAGTC	TTGCTAAATG	TGATACTGTT	GACATCTCTC	AGAATGGCCA	GAAGTGCAAT	2700
	TAACCTCTTA	GGTGGCAAGG	AGGCAGGAAG	TGCCTCTTTA	GTCTTTACAT	TTCTAATAGC	2760
	CTTGGGTTTA	TTTGCAAAGG	AAGCTTGAAA	AATATGAGAA	AAGTTGCTTG	AAGTGCAATTA	2820
	CAGGTGTTTG	TGAAGTCACA	TAATCTACGG	GGCTAGGGCG	AGAGAGGCCA	GGGATTGTGT	2880
	CACAGTACT	TGAATTAATT	CATCCAAATG	TACTAGGTTT	ACCACACACT	TGACTACOGA	2940
	TGTGATCAAC	ACTAACAAAG	AAACAAATTC	AAGGACAACC	TGCTTTTGAG	CCAGGSCAGG	3000
15	TCTCAGACAC	CCTGCTGTGT	GCCCCGCTC	CACCTCATCC	TGCCCGAAT	GCCAGTGCTC	3060
	CGAGCTCAGA	CAGAGGAAGC	CCTGCGAATA	GTTCATCAG	GCTGTTTCTC	AAAGGATGTG	3120
	TGAACGGGAG	ATGATGCACT	GTGTTTTGAA	AGTTGTCAAT	TTAAAGCATT	TTAGCACAGT	3180
	TCAATGTCCA	CAGTTGATGC	AGCATCTGTA	GATTTAAAT	CCTGAAGTGT	GGGTGCGCGA	3240
	CACACCAAGT	AGGAGCTAG	TCAGGCAGTT	TGCTTAAGGA	ACTTTTGTTC	TCGTCTCTCT	3300
20	TTCTCTTAAA	TTGGGGGTAA	GGAGCGAAGG	AAGAGGAA	AGATGACTA	ACTAAATCA	3360
	TTTTTACAGC	AAAACTGCT	CAAGCCATT	TAAATTATAT	CCTCATTTTA	AAAGTTACAT	3420
	TTGCAATAT	TTCTCCCTAT	GATAATGCAG	TGATAGTGT	GCACCTTTTC	TCTCTCTCTC	3480
	TCTCTCTCAC	ACACACACAC	ACACACACAC	ACACACACAC	AGAGACACCG	CACCATCTCG	3540
	CCTGGGGCAC	TGGAACACAT	TCCTGGGGGT	CACCGATGTT	CAGAGTCACT	AGAAATTACC	3600
25	TGAGTATCTC	TGGGAGGCCT	CATGCTCTCT	GTGGGCTTTT	TACCACCACT	GTGCGAGAGA	3660
	ACACACAGAG	GAATATGTGC	TCCCTCCAA	GCCCCAAGC	CTCAGAGAAA	GGGTGTTTCT	3720
	GGTTTGTCTC	TAGCAATGCA	TCGGTCTCTG	AGGTGACACT	CTGAGTGTGT	TGAAGGGCCA	3780
	CAAGGTGCGA	GGTTAATACT	CTTGCCAGTT	TTGAAATATA	GATGCTATGG	TTCAGATTGT	3840
	TTTTAATAGA	AACTAAAGG	GGCAGGGGAA	GTGAAGGAA	AGATGGAGGT	TTTGTGCGGC	3900
30	TGATGGGGC	ATTGGAACT	TCTTTTAAA	GTCTCTCAT	GGTCTCCAGT	TTTCAGTTGG	3960
	AACTCTGTGT	TTTAAACATT	AAGGGAGACA	AAGGCTGTGT	CCATTGCGCA	AAACTTCCTT	4020
	GGCCACGAGA	CTCTAGGTGA	TGTGTGAAGC	TGGCGAGTCT	GTGGTGTGGA	GAGCAGCCAT	4080
	CTGTCTGGCC	ATTTCAGAGG	TTCTAAAGAC	ATGGCTGGAT	GGCTGCTGTA	CCAACATCAG	4140
35	CACCTTAAAT	AATGCAATG	CAACATTCTC	CCCTCTGGGC	CTTGAAATTC	CTTGCCCTTA	4200
	TCATTGTGGG	TGAAGGAGAC	ATTTCTGTCC	TGGCTTCCC	ACAGCCCCAA	CGCAGTCTGT	4260
	GTATGATTC	TGGGATCCAA	CGAGCCCTCC	TATTTTCACA	GTGTTCTGAT	TGCTCTCACA	4320
	GCCAGGCCCC	ATCGTCTGTT	CTCTGAATGC	AGCCCTGTTT	TCAACAACAG	GGAGGTGATG	4380
40	GAACCCCTCT	GTGGAACCCA	CAAGGGGAGA	AATGGGTGAT	AAAGAAATCCA	GTTCCTCAAA	4440
	ACCTTCTGGC	GCAAGCTGGG	TCCCTCTCCT	GCTGGGTGGT	GCTTTCTCTT	GCACACCACT	4500
	CCCCACAGCG	GGGGAGAGCC	AGCAACCCAA	CCAGACAGCT	CAGGTGTGTC	ATCTGATGGA	4560
	AACCACTGGC	CTCAACACAG	TGCTTTATTC	TCCGTGTTAT	TTTGTCTGTT	ACTTTGAAGC	4620
45	ATGGAATTC	TGTTTGGGG	GATCTTGGGG	CTACAGTAGT	GGGTAAACAA	ATGCCACCG	4680
	GCCAAAGGCG	CATTAAACAA	TCGTCTCTGT	CCTGAGGGGC	CCCAGCTTGC	TGGGGCTGTG	4740
	AGGGCTCTGAT	AATCCAAGGG	TCACAGTATG	GGGAGAGGTG	CACCTGCCCA	CCTGCTAACT	4800
	CTCTGGCTAGA	CACAGTGTGT	CTGCCCAGGT	GACCTGTTCA	GCAGCAGAAC	AAGCCAGGGC	4860
	CATGGGGAGC	GGGGAAGTTT	TCACCTGGAG	ATGGACACCA	AGACAATGAA	GATTTGTTGT	4920
	CCAAATAGGT	CAATAATTCT	GGGAGACTCT	TGGAATAAAC	TGAATATATT	CAGGACCAC	4980
	TCTCTCCCTC	CCCTCATCCC	ACATCTCAAA	GCAGACAATG	TAAAGAGAGA	ACATCTCACA	5040
50	CACCCAGCTC	GCCATGCTTA	CTCATCTCTG	AATTTCAAGT	GCCATCACTG	CTCTTTCTTT	5100
	CTTCTTTGTC	ATTTGAGAAA	GGATGCAGGA	GGACAATTCC	CACAGATAAT	CTGAGGAATG	5160
	CAGAAAACAC	AGGGCAGGAC	AGTTATCGAC	AATGCATTAG	AACCTTGGTA	GCATCTCTGT	5220
	TAGAGGGACT	CCACCCCTGC	TCAACAGCTT	GGCTTCCAGG	CAAGACCAAC	CACATCTGGT	5280
	CTCTGCTCTC	GGTGCCCCAC	ACACCTAAGC	GTACTCTGTA	TTGCCATAGC	ATCATGTATG	5340
55	AACCATCTTA	CGTGTAGCAC	TACGACGTTA	TGTTTGGGTA	ATGTGGGGAT	GAACCTGCATG	5400
	AGGGCTCTGAT	TAAGGATGTG	GGGAAGTGGG	CTGCGGTAC	TGTGGGCTTT	GCAAGGCCAC	5460
	CTGGAGGCTC	GTCTGTTAGC	CAGTGTGTGA	GGAGCAAGGC	TTCAGGAAGG	GCCAGCCACA	5520
	TGCCATCTTC	CCTGCGATCA	GGCAAAAAG	TGGAATTAAG	AAGTCAAAAC	TTTATATGCA	5580
	TGTGTTATGT	CCATTTTGCA	GGATGAACGT	AGTTTAAAG	AATTTTTTTT	TCTCTTCAAG	5640
60	TTGCTTTGTC	TTTTCCATCC	TCATCACAAG	COCTGTGTTG	AGTGTCTTAT	CCCTGAGCAA	5700
	TCTTTGATGA	GATGGAGATG	ATCAATTAGT	ACTTTGTGTT	CAACCTTTAT	TCCTGTAAAT	5760
	ATTTCTGTGA	AAACTAGGAG	AACAGAGATG	AGATTGACA	AAAAAAATTT	GAATTAATAA	5820
	TAACACAGTC	TTTTTAAATC	TAACATAGGA	AAGCCTTTCC	TATATTATTCT	CTTCTAGCTT	5880
	TCTCCATTGT	TAAATCAGG	AAACAGGAA	AACACAGCTT	TCTAGCAGCT	GCAAAATGGT	5940
65	TAAATGCCCC	CTACATATTT	CCATCACTTT	GAACATAGC	TTTAGCTTGG	GAATCTGAGA	6000
	TATGATCCCA	GAACATCTT	GTCTCTACTT	CGGCTGCAAA	ACCCATGGTT	TAAATCTATA	6060
	TGGTTTGTGC	ATTTTCTCAA	CTAAAAATAG	AGATGATAAT	CCGAATTCCT	CATATATTCA	6120
	CTAATCAAG	ACACTATTTT	CATACTAGAT	TCCTGAGACA	AATACTCACT	GAAGGGCTTG	6180
	TTTAAAAATA	AATTGTGTTT	TGGTCTGTTC	TTGTAGATAA	TGCCCTTCTA	TTTTAGGTAG	6240
70	AAGCTCTGGA	ATCCCTTTAT	TGTGCTGTG	CTCTTATCTG	CAAGGTGGCA	AGCAGTCTCT	6300
	TTCAGCAGAT	TTTGCCCACT	ATTCTCTGTA	GCTGAAGTTC	TTTGCATAGA	TTTGGCTTAA	6360
	GCTTGAATTA	GATCCCTGCA	AAGGCTTGCT	CTGTGATGTC	AGATGTAATT	GTAAATGTCA	6420
	GTAATCACTT	CATGAATGCT	AAATGAGAA	GTAAGTATTT	TTAAATGTGT	GTATTTCAAA	6480
	TTTGTGTTAG	TAACTCTGGA	ATTACAAGAT	TTCTATGCG	GATTTACCTT	CATCTGTGTC	6540
75	ATGTTTCCCA	AACCTGAGG	AGGGAAGGCT	CAGAGATCGA	GCTTCTCTCT	TGAGTTCTAA	6600
	CAAAATGGTG	CTTTGAGGGT	CAGCCTTTAG	GAAGGTGCG	CTTTGTGTCT	CTTTGAGCTT	6660
	TCTGTTATGT	GCCTATCCTA	ATAAATCTCT	AAACACATT			

Seq ID No: 131 Protein sequence:
Protein Accession #: NP_036204.1

```

5      1      11      21      31      41      51
|      |      |      |      |      |
MATSMGLLLL LLLLLTPGA GTGADTEAVV CVGTACYTAH SGKLSAABAQ NHCNQNGGNL 60
ATVSKSEBAQ HVQRVLAQLL RREAALTARM SKFWIGLQRE KGKCLDPSLP LKGF5WVGGG 120
BDTPYSNMWK ELRNSCISKR CVSLLLDLQ PLLENRLPKM SBGPGSPGS PGSNIEGFVC 180
KPSFKGMCRP LALGQPGQVT YTTPTQTTSS SLEAVPFASA ANVACGEGDK DETQSHYFLC 240
KEKAPDVFDW GSSGPLCVSP KYGCNPNNGG CHQDCFEGGD GSFLOCRPG FRILLDLVTC 300
ASRNPCSSSP CRGGATCVLG PHGKNYTCRC PQGYQLDSSQ LDCVDVDECQ DSPCAQBCVN 360
TPGGPRCECW VGYRPGPGG GACQDVDECA LGRSPCAQGC TNDGSHFCS CREGYVLAGE 420
DGTQQQDVDE CVRPGGPLCD SLCPNTQGSF HCGCLPGWVL APNGVSCIMG PVSLGPPSGP 480
PDERDKGEKE GSTVPRATA SPTRGPEGTP KATPTTSRPS LSSDAPITSA PLKMLAPSGS 540
SGVWREFSIH HATAASGPQE PAGGSSSVAT QNNDGTDGQK LLLFYILGTV VAILLLALA 600
LGLLVYKRKR AKREKKEKK PQNAADSYSW VPERAESRAM ENQYSPTPGT DC

```

Seq ID NO: 132 DNA sequence
Nucleic Acid Accession #: NM_000963.1
Coding sequence: 135-1949 (underlined sequences correspond to start and stop codons)

```

25      1      11      21      31      41      51
|      |      |      |      |      |
CAATTGTTCAT ACGACTTGCA GTGAGCGTCA GGAGCACGTC CAGGAAGTCC TCAGCAGCGC 60
CTCCTTCAGC TCCACAGCCA GACGCGCTCA GACAGCAAAG CCTACCCCGG CGCCGCGCCC 120
TGCCCGCGCG TCGGATGCTC GCCCGCGCCC TGCTGCTGTG CGCGGTCTGT GCGCTCAGCC 180
ATACAGCAAA TCCTTGCTGT TCCCACCCAT GTCAAAACCG AGGTGTATGT ATGAGTGTGG 240
GATTGTGACD GTATAAGTGC GATTGTACCC GGACAGGATT CTATGAGAAA AACTGCTCAA 300
CACCGGAATT TTTGACAAGA ATAAATATTAT TTCTGAAACC CACTCCAAAC ACAGTGCAC 360
ACATACTTAC CCACTTCAAG GAATTTTGA ACGTTGTGAA TAACATTCCC TTCCTTGGAA 420
ACATAATTAT GAGTTATGTC TTGACATCCA GATCACATT GATTGACAGT CCACCAACTT 480
ACAAATCTGA CTATGGCTAC AAAAGCTGGG AAGCCTCTCT TAACCTCTCC TATTATAC 540
GAGCCCTTCC TCCTGTGCTC GATGATTGCC CGACTCCCTT GGGTGTCAAA GGTAAAAAGC 600
AGGTCCTCTG TCAAAATGAG ATTGTGGAAA AATTGCTTCT AAGAAGAAA GTTCATCCCTG 660
ATCCCGAGGG CTCAAACATG ATGTTTGCAT TCTTTGCCCA GCACCTCACG CATCAGTTT 720
TCAAGACAGA TCATAAGCGA GGGCCAGCTT TCACCAACGG GCTGGGCCAT GGGGTGGACT 780
TAAATCATAT TTACGTGAA ACTCTGGCTA GACAGCGTAA ACTGCGCCTT TTCAAGGATG 840
GAAATATGAA ATATCAGATA ATTGATGGAG AGATGTATCC TCCACAGCTC AAAGATAC 900
AGGCGAGAGT GATCTACCTC CTTCAAGTCC CTGAGCATCT ACGTTTGTCT GTGGGGCAGG 960
AGGTCCTTGG TCTGTGCTCT GGTCTGATGA TGTATGCCAC AATCTGGCTG CGGGAACACA 1020
ACAGAGTATG CGATGTGCTT AAACAGGAGC ATCCTGAATG GGGTGATGAG CAGTTGTGTC 1080
AGACAGAGAG GCTAATACCT ATAGGAGAGA CTATTAAGAT TGTGATTGAA GATTATGTGC 1140
AAACACTTGA TGGCTATCAC TTCAAACCTA AATTGAGACC AGAACTACTT TTCAACAAAC 1200
AATTCCAGTA CAAAATCGT ATTGCTGCTG AATTTAACAC CCTCTATCAC TGGCATCCCC 1260
TTCTGCTGGA CACCTTTCAA ATTCATGACC AGAAATACAA CTATCAACAG TTTATCTACA 1320
ACAACTCTAT ATTGCTGAA CATGGAATTA CCCAGTTTGT TGAATCATTC ACCAGGCATA 1380
TTGCTGGCAG GGTGTGCTGT GTTAGGAATG TTCCACCGCG AGTACAGAAA GTATCAGAGG 1440
CTTCAATTGA CCAGAGCAGG CAGATGAAAT ACCAGTCTTT TAATGAGTAC GCAGAACGCT 1500
TTATGCTGAA GCCCTATGAA TCATTGGAAG AACTTACAGG AGAAAGGAAA ATGCTGCAG 1560
AGTTGGAAGC ACTCTATGGT GACATCGATG CTGTGGAGCT GTATCTGCC CTCTGTGTAG 1620
AAAAGCCTCG GCCAGATGCC ATCTTTGGTG AAACCATGGT AGAAGTTGGA GCACCATCT 1680
CCTTGAAAGG ACTTATGGGT AATGTTATAT GTTCTGCTGC CTACTGGAAG CCAAGCACTT 1740
TGGTGGAGAG AGTGGGTTTT CAAATCATCA ACATGCGCTC AATTCACTCT CTCATCTGCA 1800
ATAACGTGAA GGGCTGTCCC TTTACTTCAT TCAGTGTTC AGATCCAGAG CTCATTAAAA 1860
CAGTCACCAT CAATGCAAGT TCTTCCGCT CCGGACTAGA TGATATCAAT CCCACAGTAC 1920
TACTAAAAGA ACGTTGAGCT GAAGCTGAGA AGTCTAATGA TCATATTTAT TTATTATAT 1980
GAACCATGTC TATTAATTTA ATTATTAAAT AATATTATA TTAACCTCCT TATGTTACTT 2040
AACTCTTCT GTACAGAGAG TCAGTACTCC TGTGTGGAG AAAGGAGTCA TACTTGTGAA 2100
GACTTTTATG TCACTACTCT AAAGATTG TGTTGCTGT TAAGTTTGA AAACAGTTTT 2160
TATTCTGTTT TATAAACCAG AGAGAAATGA GTTTTGAAGT CTTTTACTT GAATTTCAC 2220
TTATATTATA AGAACGAAAG TAAAGATGTT TGAATACTTA AACACTATCA CAAGATGGCA 2280
AAATGCTGAA AGTTTTTACA CTGTGATGT TTCCAATGCA TCTTCCATGA TGCATTAGAA 2340
GTAACTAATG TTTGAAATTT TAAAGTACTT TTGGTTATTT TTCTGTCTAT AAACAAAAAC 2400
AGGTATCAGT GCATTATTAA ATGAATATTT AAATTAGACA TTACCAGTAA TTTCACTGCT 2460
ACTTTTAAAA ATCAGCAATG AAACAATAAT TTGAAATTT TAAATTCATA GGGTAGAATC 2520
ACCTGTAAAA GCTTGTTTGA TTTCTTAAAG TTATTAAACT TGTACATATA CCAAAAAGAA 2580
GCTGTCTTGG ATTTAAATCT GTAAAAATCAG ATGAAATTTT ACTACAATTG CTGTGTAAAA 2640
TACTTATATA GTGATGTTCC TTTTTCACCA AGAGTATAAA CCTTTTATG GTGACTGTTA 2700
AACTTCTCTT TTAATCAAAA ATGCCAAAT TATTAAAGGT GTGGAGCCAC TGCACTGTTA 2760
TACTTAAAAA AGAATATTTT GTTGAGATAT TCCAGATTT GTTTATATGG CTGGTAACAT 2820
GTAAATCTA TATCAGAAA AGGCTTACC TTTAAATAA GCAATACAA AGAAGAAAAC 2880
CAAAATATTG TTCAAAATTA GGTTTAAACT TTTGAAGCAA ACTTTTTTTT ATCCTGTGTC 2940

```

5 ACTGCAGGCC TGGTACTCAG ATTTTGTCTAT GAGGTTAATG AAGTACCAAG CTGTGCTTGA 3000
 ATAACGATAT GTTTTCTCAG ATTTTCTGTT GTACAGTTTA ATTTAGCAAT CCATATCACA 3060
 TTGCAAAAGT AGCAATGACC TCATAAAATA CCTCTTCAAA ATGCTTAAAT TCATTTTACA 3120
 CATTATTTT ATCTCAGTCT TGAAGCCAAT TCAGTAGGTT CATTTGGAATC AAGCCTGGCT 3180
 ACCTGCAATG TGTTCCTTTT CTTTCTTCTT TTTAGCCATT TTGCTAAGAG ACACAGTCTT 3240
 CTCATCACTT CGTTTCTCCT ATTTTGTGTT ACTAGTTTTA AGATCAGAGT TCATTTTCTT 3300
 TGGACTCTGC CTATATTTTC TTACTCTGAC TTTTGAAGT TTTGAGTAA ACCTCAGCTC 3360
 AGGACTCTGT TTTAGCTCCT CTTAAGAAGA TTAAAAGAGA AAAAAAAGG CCCTTTTAAA 3420
 AATAGTATAC ACTTATTTTA AGTGAAGAAG AGAGAATTTT ATTTATAGCT AATTTTAGCT 3480
 10 ATCTGTAAAC AAGATGGATG CAAAGAGGCT AGTGCTCAG AGAGAACTGT ACGGGGTTTG 3540
 TGACTGGAAA AAGTTACGTT CCCATCTCAA TTAATGCCCT TTCTTATTTA AAAACAAAAC 3600
 CAAATGATAT CTAAGTAGTT CTCAGCAATA ATAATAATGA CGATAATACT TCTTTTCCAC 3660
 ATCTCATTTG CACTGACATT TAATGCTACT GTATATTACT TAATTATTTG AAGATTATTA 3720
 TTTATGCTCT ATTAGGACAC TATGTTTATA AACTGTGTTT AAGCCTACAA TCATTGATTT 3780
 15 TTTTGTGTTA TGTCACAATC AGTATAITTT CTTGGGGTTT ACCTCTCTGA ATATTATGTA 3840
 AACATTCCAA AGAAATGATT GTATTAAGAT TTGTGAATAA ATTTTGTAGA ATCTGATTGG 3900
 CATTATGAGA TATTTAAGGT TGAATGTTTG TCCTTAGGAT AGGCCTATGT GCTAGCCAC 3960
 AAGATATATT GTCTCATTAG CCGTAATGTT CCATAAGACT GACCTTTTAA AATGTTTGA 4020
 20 GGGATCTGTG GATGCTTGTG TAATTTGTTT AGCCACAATT TATTGAGAAA ATATTCTGTG 4080
 TCAAGCCTG TGGGTTTTAA TATTTTAAA TCAACGCTG ATTACAGATA ATAGTATTTA 4140
 TATAAATTAAT TGAAAAAAAT TTTCTTTTGG GAAGAGGGAG AAAATGAAT AAATATCATT 4200
 AAGATTAATC CAGGAGAAAT TTTCTTACAA TTTTACGTTT AGAATGTTTA AGGTTAAGAA 4260
 AGAATAGTAT AATATGCTTG TATAAACAC TGTCACATGT TTTTTTTAAA AAAAAAATCT 4320
 25 GATTTGTGAT TAACATTGAT CTGCTGACAA AACCTGGGAA TTTGGGTTGT GTATGCGAAT 4380
 GTTCAGTGC CTCAGACAAA TGTGTATTTA ACTTATGTAA AAGATAAGTC TGGAAATAA 4440
 TGTCTGTTTA TTTTGTACT ATTTA

Seq ID No: 133 Protein sequence
 Protein Accession #: NP_000954.1

30
 1 11 21 31 41 51
 35 MLARALLLCA VLALSHTANP CSHPCQNRG VMSVGFDPQY KCDCTRTGFI GENCSPEFL 60
 TRIKLFKPT PNTVHYILTH PKGFNVNVN IPFLRMAIMS YVLTSRSHLI DSPPTYNADY 120
 GYKWEAFSPN LSYTRALPP VFDDCPTFLG VKGKQLPDS NEIVEKLLLR RKFPDPQGS 180
 NMFAFFAQH FTHQFFKTDH KRGPFTNGL GHGVLDLHIY GETLARQRKL RLFPKDGKMY 240
 QIIDGMEVFP TVKDTQAEML YPPQVPEHLR FAVGQEVFGL VEGLMYATI WLEHNEVCD 300
 40 VLKQEHPEWG DEQLFQTSRL ILIGETIKIV IEDYVQHLSS YHFKLKDPDE LLENEQFOYQ 360
 NRIAAEFNL YHWHLLPDT FOIHQDKVNY QOFIYNSIL LEHGIQFVE SFTRQIAGRV 420
 AGGNVFPVAV QKVSQSIDQ SRQMKYQSPN EYRKRFLMKP YESFEELTGE KMSASLEAL 480
 YGDIDAVELY PALLVEKPRP DAIFGETWVE VGAPFSLKGL MGNVICSPAY WKPSTFGGEV 540
 GFQINTIASI QSLICNVKG CPFTSFSVPD FELIKVTIIN ASSSRSGLED INPTVLKER 600
 STEL

Seq ID NO: 134 DNA sequence
 Nucleic Acid Accession #: XM_059648.1
 Coding sequence: 35-664 (underlined sequences correspond to start and stop codons)

50
 1 11 21 31 41 51
 55 AGGCTGCTGA GACTTCCCTC TAGAATCCTC CAACATGAGG CCTCTTGCGC CTTACCCGCT 60
 AAAATGTTCG GGGCCACAGG CAAAGGTATT TGCAAGTTTG CTGTCTATAG TTCTATGCAC 120
 AGTAACGCTA TTCTTCTCAC AACTAAAATT CCTCAAACCT AAAATCAACA GCTTTTATGC 180
 CTTTGAAGTG AAGGATGCAA AAGGAAGAAC TGTTTCTCTG GAAAGATATA AAGGCAAGT 240
 TTCCTAGTGT GTAAACGTGG CCACTGACTG CCAACTCACA GACAGAAATT ACTTAGGGCT 300
 GAAGGAACGT CACAAAGAGT TTGGACCATC CCACTTCAGC GTGTGTGGCT TTCCCTGCAA 360
 60 TCAGTTTGA GAATCGGAGC CCOCGCCAAG CAAGGAAGTA GAATCTTTTG CAAGAAAAAA 420
 CTAGCGAGTA ACTTTCCCAA TCTTCCACAA GATTAGATT CTAGGATCTG AAGGAGAAC 480
 TGCAATTAGA TTTCTGTTG ATCTTCAAAA GAAGGAACCA AGGTGGAATT TTTGGAAGTA 540
 TCTTGTCAAC CCGTAGGGTC AAGTTGTGAA GTTCTGGAAG CCAGAGGAGC CCATGTGAAT 600
 CATCAGGCCCT GACATAGCAG CTCTGTTIAG ACAAGTGATC ATAAAAAGA AAGAGGATCT 660
 65 ATGAGAAATC CATTGCGTTT CTAATAGAAC AGAGAAATGT CTCATGAGG GTTTGGTCTC 720
 ATTTTAAACA TTTTTTTTT GGAGACAGTG TCTCACTCTG TCACCCAGGC TGGAGTGCAG 780
 TAGTGGGTTC TCAGCTCATT GCAACCTCTG CCTTTTAAA CATGCTATTA AATGTGGCAA 840
 TGAAGGATTT TTTTTTAATG TATCTTGCT ATTAAGTGGT AATGAATGTT CCCAGGATGA 900
 GGTGTATACC CAAAGCAAAA ATCAAGATA GCCAAGAAAT CAACATGAAA TATATTAACT 960
 70 ACTTCTCTG ACCATACATA AGAATTCAGA ATACACAGTG ACCAATGTGC CTCAATATCT 1020
 TATTTGTTCA CTTGACATTT TCTAGGACTG TACTTGATGA AATGTCCAAC ACACTAGACC 1080
 ACTCTTTGGA TTCAAGAGCA CTGTGTATGA CTGAAATTTT TGGAAATACT GTAATGGTT 1140
 ATGTTAAATG AATAAAACAC AAATGTTGAA AAATGTAAAA TATATATACA TAGATTCAA 1200
 TCTTATATA TGTATGCTTG TTTTGTGTAC AGGATTTTGT TTTTCTTTT TAAGTACAG 1260
 75 TCTCTAGTGT TTTACTATAA CTGCTACTAT GTATGTAACT GACATATATA AATAGTCATT 1320
 TATAAATGAC CGTATTATAA CA

Seq ID No: 135 Protein sequence
 Protein Accession #: XP_059648.1

```

5      1      11      21      31      41      51
  MEPLAAYPLK CSQPRAKVFA VLLSIVLCIV TLPLLLQLKPL KPKNISFYAF EVKDAKGRIV 60
  SLEKYKQKVS LVVNVASDCQ LTRDNYLGLK SLHKSPQPSH PSVLAPPCNQ PGSEBPRPSK 120
  EVESFARKNY GVTFPIPHKI KILGSEGBPA FRFLVDSKKL BFRWNFWKYL VNPBGQVVKF 180
10  WKPEBPIEVI RPDIAALVRQ VIKKKKEDL

```

Seq ID No: 136 DNA sequence

Nucleic Acid Accession #: NM_003003.1

15 Coding sequence: 304-2451 (underlined sequences correspond to start and stop codons)

```

20      1      11      21      31      41      51
  CAAGTGGCCT CGCCGCGCCC CTTCCTCCCTC CGGCTCCCTCC GGCCTCCCTCC CGGGAACCGG 60
  CGGTGAGCTC ACGGTGCGGG ACGAGTGGAA CGAGACTGTC CCGCGGAGC CGCCGGTATG 120
  AGCCGCCCTC GCCACCCCGT GTCCAGAGCC CGGCTTTCTT GACAGAGCT AGACTTCGGG 180
  CTCCTTGAGG ATATTCAATT TTGTATGTTT GAATATCCTC TCACCATGTT CAGCATAAAG 240
  TACCATTTCT AATGATTATC CTCAACAAGA CAGGTGTGAG AGGGTTGCTG TTGCATTGCA 300
  ATCATGGTGC AAAAATACCA GTCCCCAGTG AGAGGTGACA AATACCCCTT TGAATTAAAT 360
  ATGGCTGGCT ATGAAGAAGG GTTCCCTACA TGTCTTTTGA TTCCGATGTT CGTGGGCAGT 420
  GACACTGTGA GTGAATTCAA GAGCGAAGAT GGGGCTATTG ATGTCAATTGA AAGGCGCTGC 480
  AAGCTGGATG TAGATGCACC CAGACTGCTG AAGAAGATTG CAGGAGTTGA TTATGTTTAT 540
  TTTGTCCAGA AAAATCACT GAATTCCTGG GAACGTACTT TGACATTGA GCCTTATAAT 600
  GAAACGTTT CCAATCGGGT CATCATTAAT GAGCATTGCT GCTACACCGT TCACCCGTAA 660
  AATGAAGATT GGACCTGTTT TGAACAGTCT GCAAGTTTAG ATATTAAATC TTCTTTGGT 720
  TTGAAAGTA CAGTGGAAAA AATTGTCAATG AAACAATATA CCAGCAACAT TAAAAAAGGA 780
  AAGGAATCA TCGAATACTA CCTTCGCCAA TTGAAGAAG AAGGCATAAC CTTGTGCC 840
  CATTGAGTTC CGCCTTCCAT CACGCCCTCT TCAGAGACAT CTTTCATCAT CTCCAAGAAA 900
  CAAGCAGGT CATGGCCGT GTCATCCCA GAAGCTGCC CCAAGGAGGG GCTGAGTGGT 960
  GATGCCCTCA GCAGCCCGAG TGCACTGAG CCGCTGGTGG GCACCCCTGA CGACAACTA 1020
  GATCCGAGC ACATCAAGAG ATACCTGGGC GATTGACTC CGCTGCAGGA GAGCTGCCTC 1080
  ATTAGACTCT GCCAGTGGCT CCAGGAGACC CACAAGGCGA AAATTCCAA AGATGAGCAT 1140
  ATTCTCGGT TCTCCGTGC ACGGATTTT AATATTGACA AAGCCAGAGA GATCATGTGT 1200
  CAGTCTTGA CGTGAGAAA GCAGCATCAG GTAGACTACA TTCTTGAAC CTGAGCCCTC 1260
  CCTCAGGCTC TTCAGATTA CTACGCGGGA GGCTGGCATC ATCAGCAAA AGATGGCGG 1320
  CCGCTCTAGC TGCTCAGGCT GGGCAGATG GACCCAAAG GCTTGGTAG AGCGCTCGG 1380
  GAGGAGCGCC TGCTGAGATA CGTTCTCTCC GTAAATGAAG AACGGCTAAG CGGATGCGAA 1440
  GAGATACAA AAGTCTTTGG TCGGCTATC AGCTCATGGA CCGCTGGTGT GGACTTGGAA 1500
  GGGCTGAACA TGCGCCACT GTGAGAGCTT GGTGTGAAG CGCTGCTCGG GATCATCGAG 1560
  GTGTGGAGG CCAACTACCC TGAGACACTG GCGCGCTTC TCATCTTGG GCGGCCAGG 1620
  GTATTCTCTG TGCTCTGGAC GCTGGTTAGT CCGTTCAATT ATGACAAAC CAGAAGGAG 1680
  TTCTCATTT ATGCAGGAAA TGACTACCAG GGTCTGGAG GCGTGTCTGA TTACATCGAC 1740
  AAGAGATTA TTCCAGATT CTGTAGTGGG GAGTGATGTT GCGAAGTGCC AGAGGGTGA 1800
  CTGTCCCCA AATCTCTGTA CCGGACTGCA GAGGAGCTGG AGAAGGAAGA CTTGAAGCTC 1860
  TGAGCTAGAA CATCTACCA GTCTGCAAGC GTCTTCAAAG GAGCCGCCA TGAGATTCTC 1920
  ATTCAGATTG TGGATGGCTC GTCAGTCATC ACTTGGGATT TCGACGTGTG CAAAGGGGAC 1980
  ATTGTGTTTA ACATCTATCA CTCGAAGAGG TCGCCACAC CACCCAAAAA GGACTCCCTG 2040
  GGAGCCCA CAATCACTC TCCGGGTGGG AACAAATGTC AGCTCATAGA CAAAGTCTGG 2100
  CAGCTGGGCC GCGACTACAG CATGTGGAG TCGCCTCTGA TCTGCAAGA AGGAGAAAGC 2160
  GTGAGCGTT CCAATGTGAC CAGGTGGCGG GGCTTCTACA TCTGCAAGT GAAATTCAC 2220
  AGCATGCTTG CGTGCGCGC CAGCAGCCTT CCGCGGTGG AGCAGTGCT TGCGTCCCTG 2280
  CAGGTCTCTT CGCACAAGTG TAAAGTGATG TACTACACCG AGGTGATCGG CTCGGAGGAT 2340
  TTCAGAGTT CCAATGAGAG CCTGGAGTCC AGCCACAGCG GCTTCTCCA GCTGAGTGC 2400
  GCCACACCT CTTCCAGCCA GTCCCACTCC AGCTCCATGA TCTCCAGGTA GTCCGCGCT 2460
  GCGTGCACCT AGTGTGAGA GGGGACGGCC GCGCTCTCT GAGCAGCAGC TGCACCGGCC 2520
  CACCCAGCGG CGACATTGTA CAGACTCCTC TCACCTTAG ATAGCAAATA GCTCTCAGAT 2580
  GGTAAACGTA GTCGTTGAT CCCAAAACTA CTTGGCAGG TAGTTTTAAC TCTGATCCTA 2640
  ACTTAACCTA ATAGCCATAG ATTTGTGATA CGTTGTGAC AAAATCCAAC CAGAGCGCAA 2700
  GGGCTCTCTT GAAAGAAAAG TAGTTTCTGT ACCAATTAAA GGATGACGTT GGTCTCAGAT 2760
  ATTGATGCAA AAAAATTTTC CAACGAACTC CGCATTTGCC ATTAGTGAAT GAATTCCTGT 2820
  GACATCTCTC AGAGATGGCC CCTCTCTACC TGGGACGGAA GCTGCCAGCT CGCTTCCCC 2880
  AAGTGCCTC ATGGCCGCA CGCGCCTCA CGGCCCCAT GCTTCCCGCC AGTCAAGATG 2940
  GTCTGTGAC TTAGGGCCAG CCTTGGAGT CTTATCTCT TGAGGATTCA GAGGTTCGCT 3000
  GCGGAGTACC TTGTCCAGG GCCAGACACA CCCACACCAC CCACTGTCTG CAGTGGGGCC 3060
  GGGGGCTCAG GAGGGCTCT CAGGAGCTCC TGGTGAATCC AGGAAATATG TGCCATCGTT 3120
  AAAATTAAT TCTCTTTCC ATCTTTTGA TACTTTTAG AGCAGGATT 3180
  TTCTGTATGT GAACTTGGGT GGGGGGGTTC TTCCGTTTC CTTCGTCG TGCGCCCTCT 3240
  CACCTGCACT CAGCTCCAG CCCAGTGTAG GCCATCTCT CTGTGCCCTC TGGAGGCTCA 3300
  TTGTCTCAGA GCCAGACAG TTCCAGCCAC TAGGAGGCG TCTTGAAGCC AGCAAGTCG 3360
  ATTTGCCACT TGACACTGTC CATGGGGTTT TATTAGTAG TAAGCAGCAG CTCTGCATC 3420

```

5 CACTTCAGGG TGGCGTGTGG CATGTAGGAG TCCTGCTCTT TTGTACATGG GAATTGTGGA 3480
CTCATGCGTG TGTGTGTGTG CATGTGCTGT GTGTGTGCAT GTGTGCATGA CGTGGGGGT 3540
GCTGGGGGGA CGGGGTGAGT GGAACCTTAG TTGTAGTAAT GAAGGAATCT TCACAGAAGC 3600
AAATCAGAAT ATGGGATTGG TTGCGCTTTT ACATTTTGTT TAATTCCTGA TTTTAAAGCC 3660
TGCTCTATCT GTACAGGCC CTATTATTTT CAGCTTTTGA TGGGAAAGC AGGTTATTGT 3720
AGAATCTGTC CAGAASTTGC ATAGGGGATG GCCTCCACGA TAAGGACATG CACACGCTGT 3780
TTCTGTGTGC AGCAGAGGCC GTGTTTTTCA TGCCAAACCC CAGCGGCTG TCAACTGTGT 3840
GCGTGTAGG CATGAGATC CTGTTGTGTC GCTCTCAGCT CCGCTCTGAA GGCACGTGTGT 3900
GGGTGCTGCG TGACTGGAGA GCTGTGTGGA GGCCATGTGT GCCCGTGA CAAGATCAGGA 3960
10 GGGCGGGGGA GGGACCGAGC AGCCCTCTTG CCGGTGCGGG TCAGCCTAG TGGCTGCTG 4020
CACACTGTAG ACCTCCAGG GCCTGTGCTG TGATCACCTG CCTTGGACC ACATTGTGT 4080
TTGCTCTTAG AGATCGAGCT CCTCAGTGGT ACCTGAAGCC TTGTCTCCG GAAAGCGGG 4140
TAGGTTCTGT AGTAGGGCT AGTAGGTAGG GTTAGTAGGT AGGGCTAGTA GGTAGGGCTA 4200
TAGGTAGGAG TAGTAGGTA GGGTTCGTAG GTAGGGCTGG TAGGTAGGTT TAGTAGGTAG 4260
15 GGTAGGGCTA TAGGGTCTGT AGGTAGGGCT AGTAGGTAGG GTTAGTAGGT AGGGCTAGTA 4320
GGTAGGGCTA GTAGGTAGG TTAGTAGGTA GGGTCTGTAG GTAGGGCTGG TAGGTAGGTT 4380
TAGGTAGGAG GGTAGTAGG TAGGGTCTGT AGGTAGGGCT AGTAGGTAGG GTTAGTAGGT 4440
AGGGCTAGTA GGTAGGGCTA GTAGGTAGG TTAGTAGGTA GGGTCTGTAG GTAGGGCTGG 4500
20 GTTAGTAGGT AGGGCTAGT GGTAGTAGG TAGGGCTAGT AGGTAGGGCT AGTAGGTAGG 4560
GTAGGGCTGG TAGGTAGGT GGTAGGTAGG GTTAGTAGG GTTAGTAGGT GGTAGGTAGG 4620
GTAGGGCTGG TAGGTAGGT GGTAGGTAGG GTTAGTAGG GTTAGTAGGT GGTAGGTAGG 4680
GTAGGGCTGG TAGGTAGGT GGTAGGTAGG GTTAGTAGG GTTAGTAGGT GGTAGGTAGG 4740
GGGTCTGTAG GTAGGTAGG TAGGTAGGTT TGTAGGTAGG GTTAGTAGG GGTCTGTGTC 4800
25 TGCTTCACCC TGGTCTCTCC TGTTCCTCAA TCACAGGGC CTGAGGTAGG TCCCTGCTTT 4860
CTCTTCTCTT TCTCTCTGTG CTCAGATGGC GATTTTGCTG ACAGCTGCCA AGAAATGTCT 4920
TCACCTAACA GTCTCATGT GCCCAGAGAT GTTTTAGAA CTGTTGAAT TGCAGCCATC 4980
CCCTGCCCCC TCCAGGCTG AAGATCTGTT CTTTTAAGT TGATTCGGGA GTGGCATCTT 5040
TTTATACCCA AAGACTGTAG TGCACTCTGA AGAGCTCAA GCACATGACC GCACAAATGC 5100
TTCAGGGTT TCCCTCCGAG TAATCCAATC TCACCTCCCT TGTAAGGGAA TTCTGGGGCA 5160
30 GCTATGGTTT GAGTATGAG TTGTCATGCT GTTCTACCT TTAGTACCTT GCCACTCTTT 5220
TAAACCGCTG CTGTCATTTT CCATTCTCTA GTACTAATGA TTCTTTGATT CTCCCTCTAT 5280
TATGTCTTAA TTCACTTTCC TTCTAAATTT TGTTAATTGC ATATCAAATT CTGTAAATGT 5340
TATGTAAACA TATTACCTCA CTGGTAATA CAATACTGAT AGTCTTTAAA AGATTTTTTT 5400
35 ATTGTATACA ATAATAAATG TGAACATATT AAAG

Seq ID No: 137 Protein sequence
Protein Accession #: NP_002994.1

40 1 11 21 31 41 51
MVQKYQSPVR VYKYPFELIM AAYERRFPPTC PLIPMFVSGD TVSEPKSEDG AIHVIERRCK 60
LDVDAPRLK KLAGVDVYVP VQKNSLNSRE RTLHIEAYNE TFSNRVINE HCCYTVHPEN 120
EDWTCFQGA SLDIKSPFGP BSTVBKIAMK QTTSNIKKXK EIIIEYLRQL EEZGTFVPR 180
45 WSPSPITPSS ETSSSSSKKQ AASMAVVIPE AALKEGLSDG ALSSPSAPSP VVGTPDDKLD 240
ADHKRYLGD LTPLOESCLI RLKQWLQETH KGKIPKDEHI LRFRLARDFN IDKAREIMCQ 300
SLTWKQHQV DYILETWTFP QVLQDYIYAG WHHDKDGRF LYVLRLQMD TKGLVRALGB 360
EALLRYVLSV NEERLRCEB NTRVFGRPIS SWTCLVDLEG LAMRHLNRPV VKALLRIIEV 420
50 VEANVPETLG RLILRAPRV FFWLWTLVSP FIDNTRRKF LIYAGNDYQG PGGLLDYIDK 480
EIIIPDFLSGE CMCEVPBGGL VPKSLYRTAE ELENEDLKLW TETIYQSASV FKGAPEILI 540
QIVDASVIT WDFVCKGDI VFNHYHSKRS PQPPKDKSLG AHSITSPPGN NVQLIDKVVQ 600
LGRDYSMVES PLICKBESV QGSHVTRWPG FYILQWKFS MPACAASSLP RVDDVLASLQ 660
VSHKCKVMY YTEVIGSEDF RGSMTSLESS HSGFSQLSAA TTSSSQSHSS SMISR

55 Seq ID No: 138 DNA sequence
Nucleic Acid Accession #: NM_004181.1
Coding sequence: 32-670 (underlined sequences correspond to start and stop codons)

60 1 11 21 31 41 51
GCAGAAATAG CCTAGGGAGA TCAACCCCGA GATGCTGAAC AAAGTGCTGT CCGGCTGGG 60
GGTCCCGGC CAGTGGCGCT TCGTGGAGCT GCTGGGCTG GAAGAGGAGT CTCTGGGCTC 120
GTTGCCAGCG CCGTGGCTGG CGCTGCTGCT GCTGTTTCCC CTCACGGCCC AGCATGAGAA 180
65 CTTACAGAAA AAGCAGATTG AAGAGCTGAA GGGACAAGAA GTTAGTCTTA AAGTGTACTT 240
CATGAAGCAG ACCATTGGGA ATTCTGTGG CACAATCGGA CTTATTCAAG CAGTGGCCAA 300
TAATCAAGAC AAACCTGGAT TTGAGGATGG ATCAGTTCTG AAACAGTTTC TTTCTGAAC 360
AGAGAAATG TCCCTGAAG ACAGAGCAA ATGCTTTGAA AAGAAATGAG CCATACAGGC 420
AGCCCATGAT GCGTGGCAC AGGAAGGCCA ATGTGGGTA GATGACAGG TGAATTTCCA 480
70 TTTTATTCTG TTTAACAACG TGGATGGCCA CCTCTATGAA CTTGATGAC GAATGCCTTT 540
TCCGCTGAAC CATGGGCGCA GTTCAGAGGA CACCCTGCTG AAGGACGCTG CCAAGGTGTG 600
CAGAGAAATC ACCGAGCGTG AGCAAGGAGA AGTCCGCTTC TCTGCGGTGG CTCTCTGCAA 660
GCGACCTTAA TGCTCTGTGG GAGGAGCTTT GCTGATTTCC CCTCTTCCCT TCAACATGAA 720
75 AATATATACC CCCCATGCAG TCTAAATGTC TTCAGTACTT GTGAAACACA GCTGTCTTTC 780
TGTTCTGCAG ACACGCTTC CCTCAGCCA CACCAGGCA CTTAAGCACA AGCAGATGTC 840
ACAGCTGTCC ACTGGGCCAT TGTGTGTGTA GCTTCAGATG GTGAAGCATT CTCCCCAGTG 900

TATGTCCTGT ATCCGATATC TAACGCTTTA AATGGCTACT TTGGTTCTGT TCTGTAAGTT 960
AAGACCTTGG ATGTGGTTAT GTTGTCTTAA AGAATAAATT TTGCTGATAG TAGC

5 Seq ID No: 139 Protein sequence:
Protein Accession #: NP_004172.1

10 1 11 21 31 41 51
MLNKLRLG VAGQWRPVDV LGLLEESLGS VPAPACALLL LFPLTAQHEN FRKKQIERLK 60
GQEVSPKVVY MKQTIGNSCG TIGLIHAVAN NQDKLGFREDG SVLKQFLSET EKMSPEDRAK 120
CFEKMBAIQA AHDAVAQBGQ CRVDDKVNPH FILPNNVDGH LYELOGRMPPF PVNKGASSED 180
TLLKDAAKVC REPTEREQGE VRPSAVALCK AA

15 Seq ID No: 140 DNA sequence
Nucleic Acid Accession #: NM_000201.1
Coding sequence: 58-1656 (underlined sequences correspond to start and stop codons)

20 1 11 21 31 41 51
GGGCCCACTG CGACGCTGAG CTCTCTGCT ACTCAGAGTT GCAACCTCAG CCTCGCTATG 60
GCTCCACGCA GCCCCCGGCC CGCGCTGCCC GCATCTCTGG TCTGCTCGGG GGCTCTGTTC 120
GCAGGACCTG GCATGCCCCA GACATCTGTG TCCCTCTCAA AAGTCATCCT GCCCCGGGGA 180
GGCTCCGTGC TGTGACATG CAGCACCTCC TGTGACGAGC CCAAGTTGTT GGCATAGAG 240
ACCCCGTTGC CTAAAAAGGA GTTGTCTCTG CTTGGGAACA ACCGGAAGGT GTATGAATG 300
AGCAATGTGC AAGAAGATAG CCAACCAATG TGCTATTCAA ACTGCCCTGA TGGCAGTCA 360
30 ACAGCTAAAX CCTTCTCTAC CGTGTACTGG ACTCCAGAAC GGGTGGAACT GGCACCCCTC 420
CCCTCTTGCG AGCCAGTGGG CAAGAACCTT ACCCTACGCT GCCAGGTGGA GGTGGGGCCA 480
CCCCGGGCCA ACCCTACCGT GGTGCTGTCT CGTGGGGAGA AGGAGCTGAA ACGGGAGCCA 540
GCTGTGGGGG AGCCCGCTGA GGTCAOGACC ACCGTGCTGG TGAGGAGAGA TCACCATGGA 600
GCCAATTTCCT CGTGCCTGAC TGAACCTGAC CTGCGGCCCC AAGGGCTGGA GCTGTTTGAG 660
35 AACACCTCGG CCCCCACCA GCTCCAGACC TTTGTCTGTC CAGCGACTCC CCCACACTT 720
CTCAGCCCCC GGTCTCTAGA GGTGGACACG CAGGGGACCG TGGTCTGTTC CCTGGACGGG 780
GTGTTCCCAg TCTCGGAGGC CAGGTCACAC CTGGCACTGG GGGACCAGAG GTTGAACCCC 840
ACAGTCACCT ATGGCAACGA CTCTTCTCG GCCAAGGCTT CAGTCAGTGT GACCCGAGAG 900
40 CCGAGAGGGA CCGACGGGCT GACGTGTGCA GTAATCTGG GGAACCCAGG CAGGAGACA 960
CTGCAGACAG TGACCATCTA CAGCTTTCGG GCGCCCAACG TGATTCTGAC GAAGCCAGAG 1020
GTGCAAGAG GGAACGAGGT GACAGTGAAG TGTGAGGCCC ACCCTAGAGC CAAGGTGACG 1080
CTGAATGGGG TTCCAGCCCA GCCACTGGGC CCGAGGGCCC AGCTCTCTGT GAAGGCCACC 1140
CCAGAGGACA ACCGGCGCAG CTTCTCTGTC TCTGCAACCC TGGAGGTGGC CGGCAGCTT 1200
ATACACAGA ACCAGACCCG GGAGCTTCGT GTCTGTATG GCCCCGACT GGCAGAGAGG 1260
45 GATTGTCCGG GAACCTGGAC GTGGCCAGAA AATTCCAGC AGACTCCAAT GTGCCAGGCT 1320
TGGGGGAACC CATGCCCGA GCTCAAGTGT CTAAGGATG GCATTTCCC ACTGCCATC 1380
GGGGAATCAG TGACTGTAC TCGAGATCTT GAGGGCACCT ACCTCTGTGG GGCAGGAGC 1440
ACTCAAGGGG AGGTCAACCG CGAGGTGACC GTGAATGTG TCTCCCCCG GTATGAGATT 1500
50 GTCATCATCA CTGTGGTAGC AGCCGACATC ATATATGGCA CTGCAGGCTT CAGCAGGTAC 1560
CTCTATAACC GCCAGCGGAA GATCAAGAAA TACAGACTAC AACAGGCCA AAAAGGGACC 1620
CCCATGAACG CGAACACACA AGCCACGCTT CCTTGAACCT ATCCCGGGAC AGGCCCTCTT 1680
CTCGGCCCTT CCTATATTGG TGGCAGTGGT GCCACACTGA ACAGAGTGA AGACATATG 1740
CATGCAGCTA CACTACCGG CCTGGGACG CCGAGGACA GGGCATTTGT CTCAGTCAGA 1800
TACAAACAGCA TTTGGGGCCA TGGTACCTGC ACACCTAAAA CACTAGGCCA CGCATCTGAT 1860
55 TTAGTCTCA ATGACTAAGC CAAGAGGAAG GAGCAAGACT CAAGACATGA TTGATGGATG 1920
CTAAGTCTA GCCTGATGAG AGGGGAAGTG GTGGGGAGA CATAGCCCA CCATGAGGAC 1980
ATACACTGAG GAATACTGA AACTTGTGTC CTATTGGGTA TGCTAGGCC CACAGACTTA 2040
CAGAAGAAAT GGCCTCCAT AGACATGTGT AGCATCAAA CACAAGGCC CACACTTCTT 2100
GACGATGCCC AGCTTGGSCA CTGCTGTCTA CTGACCCCAA CCCTTGATGA TATGTATTTA 2160
60 TTCTATTGTT ATTTTACCAg CTATTATTG AGTGTCTTT ATGTAGGCTA AATGAACATA 2220
GGTCTCTGGC CTCACGGAGC TCCAGTCCA TGTACATTC AAGGTACCA GGTACAGTTG 2280
TACAGGTTGT AACTGACAGG AGAGTGCCCT GCAAAAAGAT CAAATGGGGC TGGGACTTCT 2340
CATTTGCCAA CCTGCTTTC CCCAGAAGGA GTGATTTTTC TATCGGCACA AAAGCACTAT 2400
ATGACTGTGT AATGTTTCAC AGGTTTCAGAG ATTACCCAGT GAGGCCCTAT TCCTCCCTTC 2460
65 CCCCCAAAC TGACACTTTT GTTAGCCACC TCCCCACCA CATACATTTC TGCCAGTGT 2520
CACAATGACA CTCAGCGGTC ATGTCTGGAC ATGAGTGCCC AGGGAATATG CCCAAGCTAT 2580
GCCCTGTCTCT CTGTCTCTGT TTGCAITTC CTGGGAGCTT GCATATTGTC AGCTCCAGTT 2640
TCCGTCAGTG ATCAGGTTCC TGCAAGCAGT GGGGAAGGGG GCCAAGGTAT TGGAGGACTC 2700
CCTCCAGCTT TTGGAAGGGT CATCCGCGTG TGTGTGTGTG TGTATGTGTA GACAAGCTCT 2760
70 CGCTCTGTCA CCCAGGCTGG AGTGCAGTGG TGCAATCATG GTTCACTGCA GTCTTGACCT 2820
TTTGGGCTCA AGTGTATCTC CCACCTCAGC CTCTGAGTA GCTGGAGCCA TAGGCTCACA 2880
ACACCACACC TGCCAAATTT GATTTTTTTT TTTTTTTTCA GAGACGGGGT CTGCAACAT 2940
TGCCAGACTT TCCTTTGTGT TAGTTAATAA AGCTTTCTCA ACTGCC

75 Seq ID No: 141 Protein sequence:
Protein Accession #: NP_000192.1

1 11 21 31 41 51
 5 MLQFVRAGAR AWLRPTGSQG LSSLAEBAAR ATENPEQVAS EQLPEPVLRLK VELFVPVTHRR 60
 PVQAWVESLR GPEQERVGLA DLHFDVFATA PRLDILHQA MWQKNPKRIS YAKTKTRAEV 120
 RGGGKPLAA ERHWAGPAMQ HPLSALARRR CCFWPPGPTS YYMLPKMVR ALGLKVALTV 180
 KLADQDLHM DSELEPTGDF QYLTELHAYR RNQDSVLLVD LTHEEMPQSI VEATSRLLKTF 240
 NLIPAVGLHV HSMLEKHQTLV LTLPTVAFLE DKLLMQDSRY RFLYPFSLFY SDFPRPLPHA 300
 TQGPATPFH C

Seq ID NO: 142 DNA sequence

Nucleic Acid Accession #: NM_000270.1

Coding sequence: 110-979 (underlined sequences correspond to start and stop codons)

15 1 11 21 31 41 51
 AACTGTGCGA ACCAGACCCG GCAGCCTTGC TCAGTTCAGC ATAGCGGAGC GGATCCGATC 60
 AGATCGGAGC ACACCGGAGC AGGCTCATCG AGAAGGCGTC TGGCAGACCA TGGAGAACCG 120
 20 ATACACCTAT GAAGATTATA AGAACACTGC AGAATGGCTT CTGTCTCATA CTAAGCACCG 180
 ACCCTCAAGTT GCAATAATCT GTGGTTCTGG ATTAGGAGGT CTGACTGATA AATTAAGTCA 240
 GGCCAGAGAT TTGTACTACA GTGAAATCCC CAACCTTCCT CGAAGTACAG TGCCAGGTCA 300
 TGCTGCGCGA CTGGTGTGTT GGTTCCTGAA TGGCAGGCCC TGTGTGATGA TGCAGGCGAG 360
 GTTCCACATG TATGAAGGGT ACCCACTCTG GAAGGTGACA TTCCAGTGA GGGTTTTCCT 420
 25 CCTTCTGGGT GTGACACCCC TGGTAGTCAC CAATGCAGCA GGAGGGCTGA ACCCCAAGTT 480
 TGAGGTTGGA GATATCATGC TGATCCGTGA CCATATCAAC CTACTCGTT TCAGTGGTCA 540
 GAACCCCTCT AGAGGGCCCA ATGATGAAAG GTTTGGAGAT CGTTTCCTGT CCATGCTGTA 600
 TGCCACGAC CGGACTATGA GGCAGAGGGC TCTCAGTACC TGGAAACAAA TGGGGGAGCA 660
 ACGTAGCTCA CAGGAAGGCA CCTATGTGAT GGTGGCAGGC CCCAGCTTTG AGACTGTGGC 720
 30 AGAATGTGCT GTGCTGCAGA AGCTGGGAGC AGACGCTGTT GGCATGAGTA CAGTACCAGA 780
 AGTATCGTCT GCACGGCATG CTGGACTTCG AGTCCTTTGG TTCTCACTCA TCACTAACAA 840
 GGTATCATCG GATTATGAAA GCCTGGAGAA GGCCAACCAT GAAGAAGTCT TAGCAGCTGG 900
 CAAACAAGCT GCACAGAAAT TGAACACGTT TGCTCCATT CTTATGGCCA GCATTCCACT 960
 CCTGTACAAA GCGAGTTGAC CTGCTTGGG ATCTCCACA CAGAACCCAA 1020
 35 GTAGCTGCTA CCTTCTTTGG CCCCTTGCTG GAGTCATGTG CCTCTGTCTT TAGGTTGTAG 1080
 CAGAAAGGAA AGAGTTCTCT TCCTTCACCT TTCCCACTTT CTCTTACCAG ACCCTTCTGG 1140
 TGCCAGATTC TCTTCTCAA GCTGGGATTA CAGGTGTGAG CATAGTGAGA CCTTGGGCTT 1200
 ACAAAATAAA GCTGTTCTCA TTCTGTCTCT TTCTTACACA AGAGCTGGAG CCGTGGCCCT 1260
 40 ACCACACATC TGTGGAGATG CCCAGGATTG GACTCGGGCC TTAGAAGCTT GCATAGCAGC 1320
 TGTACTAGC TCTTTGAGAT AATACATTCC GAGGGGCTCA GTTCTGCCTT ATCTAAATCA 1380
 CCAGAGACCA AACAGGACT AATCCAATAC CTCTTGGG

Seq ID No: 143 Protein sequence

Protein Accession #: NP_000261.1

45 1 11 21 31 41 51
 MENGYYTDED KNTAEWLLSH TKHRPQVAII CGSGLGLTD KLTQAQIFDY SEIPNFFPRST 60
 50 VPGHAGRLVF GFLNGRACVM MQGRPHMYEG YPLMKVTFPV RVPHLLGVDV LVTVNAAGGL 120
 NPKFEVDIM LIRDHINLPG FSGQNFRLRP NDERPGRDRP AMSDAYDRTM RQRLSTWQK 180
 MGEQRELEQEG TYVMVAGPSF ETVABCRVLQ KLQADAVGMS TVPEVIVARH OGLRVFGPSL 240
 ITNKVIMDYB SLEKANHEEV LAAGKQAAQK LEQPVSIAMA SIPLFDKAS

Seq ID NO: 144 DNA sequence

Nucleic Acid Accession #: NM_015577.1

Coding sequence: 112-3054 (underlined sequences correspond to start and stop codons)

60 1 11 21 31 41 51
 GAAGCGGCGG GCGGGGTGGA GCAGCCAGCT GGGTCCGGGG AGCGCCGCGG CCGCCTCGAT 60
 GGGGTGTGTA AAGTCTCCT CTAGAGCTTT GGAAGGCTGA ATGCACTAAA CATGAAGAGC 120
 TTGAAGCGGA AGTTCAGGAA GAGTGACACC AATGAGTGA ACAAGAATGA TGACCGGCTA 180
 65 CTGACGCGCG TGGAGAATGG AGATGCGGAG AAGGTGGCCT CACTGCTCGG CAAGAAGGGG 240
 GGCACTGCCA CCAACACGA CAGTGAGGGC AAGACCGCTT TCCATCTTGC TGCTGCAAAA 300
 GACACAGTGG AATGCCCTCAG GGTCAATGAT ACACATGGTG TGGATGTGAC AGCCCAAGAT 360
 ACTACCGGAC ACAGCGGCTT ACATCTCGCA GCCAAGAACA GCCACCTAGA ATGCATCAGG 420
 70 AGGCTGCTTC AGTCTAATG CCCAGCCGAA AGTGTGACCA GCTCTGGGAA AACAGCTTTA 480
 CATTATGCAG CGGCTCAGGG CTGCCTTCAA GCTGTGCAGA TTCTCTGCGA ACACAAGAGC 540
 CCGATAAACG TCAAGATTT GATGGGAAT ATACCGCTGC TTCTTGTCTG ACAAAATGGT 600
 CACAGTGAGA TCTGTCACTT TCTCCTGGAT CATGGAGCAG ATGTCAATTC CAGGAACAAA 660
 ATGTGAAGAA CTGCTCTCAT GCTGGCCTGT GAGATTGGCA GCTCTAACGC TGTGGAAGCC 720
 75 TTAATTAAAA AGGGTGACGA CCTAAACCTT GTAGATTCTC TTGATACAAA TGCCTTACAT 780
 TATTCCAAAC TCTCAGAAAA TGCAAGAAAT CAAAGCCTTC TATTATCAAA AATCTCTCAG 840
 GATGCTGATT TAAAGACCCC AACAAAACCA AAGCAGCATG ACCAAGTCTC TAAAAAAGC 900

TCAGAAAGAA GTGGAACCTC AAAAACACGC AAAGCTCCAC CACCTCCTAT CAGTCTACC 960
 CAGTGTAGTG ATGTCTCTTC CCAAGATCA ATAACTTCCA CTCCACTATC GGAAGAGGAA 1020
 TOGTAITTTT TTGCTGAACC ACCCTTCAAG GCTGAGATCA GTTCTATACO AGAAACAAA 1080
 GACAGACTAA GTGACAGTAC TACAGGTGCT GATAGCTTAT TGGATATAAG TTCTGAAGCT 1140
 GACCAACAAG ATCTTCTCTC TCTATTGCAA GCAAAAGTTG CTTCCTTAC CTTCACAAAT 1200
 AAGGAGTTAC AAGATAAATT ACAGGCCAAA TCACCAAGG AGGCGGAAGC AGACCTAAGC 1260
 TTTGACTCAT ACCATTCCAC CCAAACTGAC TTGGGCCCAT CCGTGGGAAA ACCTGGTGAA 1320
 ACCTCTCCCC CAGACTCCAA ATCATCTCCA TCTGTCTTAA TACATCTTT AGGTAAATCC 1380
 ACTACTGACA ATGATGTGAG AATTGAGCAA CTGCAAGAGA TTTTGCAAGA TCTACAGAA 1440
 AGATTAGAGA GCTCTGAAGC AGAGAGAAA CAGCTACAGG TCGAACTCCA ATCCGAAAG 1500
 GCAGAACTGG TATGCTTAAA CAACACTGAG ATTTGAGAGA ACAGCTCTGA CCTCAGCCAG 1560
 AAATCTAAAG AAATCTAGAG CAAATACGAG GAGGCTATGA AAGAAGTCTT TAGTGTGAG 1620
 AAGCAGATGA AACTCGGTCT TGTCTCACTT GAAAGCATGG ATAATTATTC ACATTTCAC 1680
 GAGCTGAGGG TCACGGAAGA GGAATAAAT GTGCTAAAGC AGGATCTGCA GAATGCATTA 1740
 GAAGAAAGTG AAAGAAATAA AGAGAAAGTG AGAGAGTTAG AGGAAAACT GGTAGAGAGG 1800
 GAGAAAGGTA CAGTGATTA GGCACCTGTG GAAGAGTACG AGGAAATGAA AAGTTCATAT 1860
 TGCTCTGTTA TTGAGAAATAT GAATAAGGAG AAAGCATTTT TGTTTGAGAA ATACCAAGAA 1920
 GCGCAAGAGG AAATCATGAA ATTAAGAGAC ACATAAATA GTGAGATGAC ACAGGAAGCC 1980
 AOTGATGAAG CTGAGGACAT GAAAGAAAGC ATGAATAGGA TGAATAGTA ACTCAATAA 2040
 CAGGTGAGCG AGCTGTGACA GCTGTACAAA GAAAGCCAGG CTGAGCTGGA GGATTACAGG 2100
 AAGAGAAAT CTCTAGAGGA TGTACAGCT GAATATATCC ATAAAGCAGA GCATGAGAAA 2160
 CTGATGCAAT TGACAAAGCT GTCCAGGGCT AAAGCAGAG ATGCACTGTG TGAATGAG 2220
 TCTCAGATT CAAAGTGTG GAATGAGTTG ACCCAGCTCA AACCACTGGT GGATGCACAA 2280
 AAAGAGAACT CTGTCTCTAT CACAGACAT TTGCAAGTGA TAACCAAGCT GCGACTGCA 2340
 GCAGAGAGG TCAGAGAAA AATAAGCAAT CTTAAGGAAC ACCTTGCAAG CAAGGAAGTG 2400
 GAAGTAGCAA AGCTGGAGAA ACAACTCTTA GAAAGAGAA CTGCTATGAC TGATGCAATG 2460
 GTACCTCGGT CTTCCTATGA AAACTCCAG TCATCCTTAG AGAGTGAAGT GAGTGTGTTG 2520
 GCATCGAAAT TAAAGGAATC TGTGAAAGAG AAAGAGAGG TCCATTGAGA GGTGTGCCAG 2580
 ATTAGAGAGT AGGTCTCACA GTTGAAAGAG GAAAGAGAAA ATATTGAGC TCTCTTGAAA 2640
 TCCAAAGAGC AAGAGTAAA TGAACCTCTG CAAAATTTCC AGCAAGCTCA GGAAGAACTT 2700
 CGAGAAATGA AAGATACGCG TGAGAGCTCT TCAAACTGCG AGGAAGATA AGATAAAAG 2760
 ATAAATGAGA GTGCGAAGGA AGTCAACAAA TTGAAGGAGG CCTTGAACAG CCTTCCAG 2820
 CTCTCTACT CAACAAGCTC ATCCAAAGAG CAGAGTCAGC AGCTGGAGGC GCTGCAAGC 2880
 CAACTCAAAC AGCTCCAGAA CCAGCTGGCG GAATGCAAGA AACACACCA GGAGTCATA 2940
 TCAATTTACA GAATGCATCT TCTGTATGCT GTGCAAGGCG AGATGATGA AGATGTCCAG 3000
 AAGTACTGGA AGCAAACTCT TACCATGTTT AAAAACAGT CTCAAAAGAA GTAAAGTGA 3060
 TTCTTTGGCA GGCACTGCC CCTTGTCTAT TGTCTTTGTC TTGATCCAG AGTGTGCGC 3120
 AGCGGCTGCC ATTTGTCTCA TTGCTGTAT GCATGTGCG CTAGCTAGC TTCTTCCCTT 3180
 TCCAAAGGTT TCTGAGGACT TCTCCAGGA GAAGACTGCC CGCTCAGAA CTGCTTAGAG 3240
 ACTTCAAAAC AGCAGAGGTT AAGTCCCTG TCATCCCTTC AGATTCCAGA GCTGGATCA 3300
 GCGTGGCCA GAGGTCTGGT CCGATGCTG CGAGGGGGGC CCCTCTCTCC ATCCCTGACT 3360
 GGCATGAGTG CTTTATCACC ACCGAGTGAT GTGCTGAGGC CTCTGCACT GAATGCTCT 3420
 TCCATTTCTG TCTCGGGCA GTGCCATTCA GCACAGGAGA GCTCTTTTTC CCTTTGGCTT 3480
 TCAATTTCAA AACATGATT AATTCTAAC TAAATTAGTA TGGCACTAGT TATGAAGTAT 3540
 CTGCTTAAAA CCTTTCATCA TGATATCTCT TGGATTTAAA AACTCTAAT CCATGTTTT 3600
 TTCCATCTG CCTTATATAT CTATCACCC TGCTTATCAA TATTCACTT GATGAGCACT 3660
 ATTAATTAAT ATATGAAGT TAAACAAA AGCAAGTTGT CCTTAAAGT TCTTTTTTA 3720
 AGTAAATGT TGACATAGT CAAATTTCT ATGCAAACT GCCTCTGCT GTTATCTGTG 3780
 AAGCTCAGGA AATCCAAACA TTTGTGTTTC AACAGGGAC AGTAACTGT GTGTTACAG 3840
 CCAAAAGAAA TGCTCATAG TTCTTAACCT CAACTTTTGT AGAAGTATTT TTTTCTCTGT 3900
 AATATTTTTA TTGGCTCATA AAGATGTTT CATATCTGAA CTCTTAAATA AGTGAATTA 3960
 CAGTAGATTA TATTACAAA ATACTTTTTA GGTAGCCATG CTGAGACTT TTTAAAAATA 4020
 TAACTTTTTC CTTAAAGTTT TCAGCTATAG CAAAAGTAG TTATGTATGC CAGACCTAAT 4080
 ATGAGCTGCC ACCAACACCC CTAGAATTT CAGCCATGGT GTCTTCAGAA TTGTAGCGCA 4140
 TTTCTGAATC TAGCAATCC TCCTTTTACC CGTTGAATGT TTGAATGCC CTGACTCTAC 4200
 CAGCGCCCAT AAATGATCTC TAGAAGGACT GTTAGTACCA ATCTGTTTT CAACTTTGAA 4260
 GCTTAAAAAC CTGATATGGT AATATTATGG TGATAGCAG AGGTCTCGGA AAAAAAATAT 4320
 TTCTGTTTCC TTTACTTTCA GGTAAAAAT GTTCTAACA CGCTTGCAAC TTCCCTTATG 4380
 CATTAATCT TGTGAGGGA GAGAGACAGA ATCTGAGCT CTCCAAAGTA TTTAACTGAA 4440
 AGTAGGGCT GCTCTGACAG GCGCCATGTC CCACAAGCT CTTGGGCTC AGTGGGTGCT 4500
 TGCTGTGCT GATGTATATG TTGATCTGTA TTGATTAAG ACCAATGACA GCAAGCAAA 4560
 AATGCTTTA AAGCTTGGTG TTACTTTTCT TAAAGTTTGT AATTATAGTT AAGCAATTC 4620
 AAAAATGCTC CAAAGAAATG TGAAGAGACC TTTTGTACA GCACTTCAGA AATAACAAA 4680
 CAGCCCCCTC TGCCCCGCCA CAGAAATGCT GCAGAGTATA TAAACTTGA GACATTTTGT 4740
 TAGGATGCTC GACGAGGTGT AGCCTTTTAT CTGTGTTCCG GATGCATATT TATTACAGT 4800
 ACTCTGTTA AATATTGAAA AGTTATATGC TGTAGTTTTT AGTATTTTGT CTTTGTAAAT 4860
 TACAGAAATT ATTGAGAGAA ATAAACTTGT TTCATTTTGC AAAAAAATA AAAAAAATA 4920
 AAAAA

70 Seq ID No: 145 Protein sequence:
 Protein Accession #: NP_056392.1

75 1 11 21 31 41 51
 | | | | |
 MKSLKAKFRK SDTNEWNKND DRLLQAVENG DAEKVASLLG KKGASATKHD SEGKTAFHLA 60

AAKGHVEICR VMITHQVDVT AQDTTGHSAI HIAAKNSHBE CIRRLIQSKC PAESVDSGGK 120
 TALHYAAAGG CLQAVQILCB HKSPINLKDL DGNIPILLAV QNGHSEICHP LLDHGADVNS 180
 RNKSGRTALM LACBIGSSNA VEALIKKGAD LNLVDSLGYN ALHYSKLSN AGIQSLLLSK 240
 ISQADLKTPT TKPKQHDQVS KISSERSGTP KTRKAPPPPI SPTQLSDVSS PRSITSTPLS 300
 GKESVFFABP PFKAEISSIR ENKDRISDST TGADSLDIS SEADQDPLLS LLQAKVASLT 360
 LHNKELQDKL QAKSPKEAEA DLSFDSYHST QTDLGPSLGK PGTSPPDSEK SSPSVLIHSL 420
 GKSTTNDVDR IQQLQBILOD LQKRLESSEA ERKQLQVELQ SRRAELVCLN NTEISENSSD 480
 LSQKLKSTQS KYERANKVEL SVQKQMKLGL VSPESMDNYS HFHELRTVEE EINVLKQDLQ 540
 NALRESERNK EKVRLEBEKL VEREKSTVIK PVVEYREMK SSYCSVIENM NKEKAFLEK 600
 YQEAQBEIMK LKDTLKSQMT QEASDEARDM KEAMGRMIDE LNKQVSELSQ LYKEAQAELE 660
 DYRRKRSLED VTAHYIKHAE HSKLMQLTNV SRARAEALS EMKSQYSKVL NBLTQLKQLV 720
 DAQKENSVDI TEHLQVITTL RTAAKEMEEK ISNLEKHLAS KEVEVAKLEK QLLSEKAAMT 780
 DAMVPRSSYE KLQSSLESSE SVLASKLKEK VKSEKRVHSE VVQIRSEVSQ VKREKEMIQT 840
 LLKSKQEVN RLQLQFQQAQ BELAEMKRYA BSSSKLEEDK DKKINEMSKE VTKLKEALNS 900
 LKQSYSTSS SKRQSQQLQA LQQQVKQLQN QLACCKKHQ EVISVYRMHL LYAVQGMDE 960
 DVQVVKQLL TMCKKQSQKK

Seq ID NO: 146 DNA sequence
 Nucleic Acid Accession #: NM_000459.1
 Coding sequence: 149-3523 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 25 CTCTCTGCT GTTCTCTCTT GCCTCTAACT TGTAAACAAG ACGTACTAGG ACGATGCTAA 60
 TGGAAAGTCA CAACCCGCTG GGTCTTTTGA AGGATCCTTG GGACCTCATG CACATTTGTG 120
 GAACTGTGAT GGAGAGATT GGGGAAGCAT GGACTCTTTA GCCAGCTTAG TTCTCTGTGG 180
 AGTCACCTGT CTCTCTTCTG GAAGTGTGGA AGGTGCCATG GACTTGATCT TGATCAATTC 240
 CCTACCTCTT GTATCTGATG CTGAACATC TCTCACCTGC ATTGCCTCTG GGTGGGCCCC 300
 CCATGAGCCC ATCACCATTAG GAAGGGACTT TGAAGCCTTA ATGAACGAGC ACCAGGATCC 360
 CTCTGGAGTT ACTCAAGATG TGACCAAGAG ATGGGCTAAA AAGATTGTGT GGAAGAGAGA 420
 AAGGCTATAT AAGATCAATG GTGCTTATTT CTGTGAAGGG CGAGTTCGAG GAGAGGCAAT 480
 CAGGATACGA ACCATCAAGA TGCCTCAACA AGCTTCTCTC CTACAGCTTA CTTTAACAT 540
 GACTGTGGAT AAGGGAGATA ACGTGAACAT ATCTTTCAAA AAGTATTGTA TTAAGAAGA 600
 AGATGCAGTG ATTTACAAAA ATGGTTCCTT CATCCATCA GTGCCCCGGC ATGAAGTACC 660
 TGAATTTCTA GAAGTACACC TGCTCATGTC TCAGCCCCAG GATGCTGGAG TGTACTCGGC 720
 CAGGTATATA GGAGGAAACC TCTTCACTC GGCCTTCAAC AGGCTGATAG TCCGGAGATG 780
 TGAAGCCAG AAGTGGGAC CTGAATGCAA CCATCTCTGT ACTGCTTGTA TGAACAATGG 840
 TGTCTGCCAT GAAGATACTG GAGAATGCAT TTGCCCTCTT GGTATTATGG GAAGGACGTG 900
 40 TGAGAAGGCT TGTGAAGTGC ACAGGTTTGG CAGAAGTGT AAGAAAGGT GCAGTGGACA 960
 AGAGGAGTGC AAGTCTTATG TGTCTGTCTT CCTGACCCCT TATGGGTGTT CCGTGTGCCA 1020
 AGGCTGGAGT GGTCTGAGT GCAATGAAGC ATGCCACCTT GGTTTTACG GGCCAGATTG 1080
 TAAGCTTAGG TGCAGCTGCA ACAATGGGGA GATGTGTGAT CGCTTCCAAG GATGTCTCTG 1140
 CTCTCCAGGA TGGCAGGGGC TCCAGTGTGA GAGAGAAGGC ATACCGAGGA TGACCCCAAA 1200
 45 GATAGTATG TTGCCAGATC ATATAGAAGT AAACAGTGGT AAATTTAATC CCAATTTCAA 1260
 AGCTTCTGCG TGGCCGCTAC CTACTAATGA AGAATGACC CTGCTGAAGC CGGATGGGAC 1320
 AGTGCTCCAT CCAAAAGACT TTAACCATAC GGAATCATTG TCAGTAGCCA TATTCACCAT 1380
 CCACCGGATC CTCCCCCTG ACTCAGGAGT TTGGGTCTGC AGTGTGAACA CAGTGGCTGG 1440
 GATGGTGGAA AAGCCCTTCA ACATTCTCTT TAAAGTCTT CCAAGCCGCC TGAATGCCCC 1500
 50 AAAAGTGAAT GACACTGGAC ATAACTTTGC TGTCAATCA ATCAGCTCTG AGCCTTACTT 1560
 TGGGGATGGA CCAATCAAT CCAAGAAGCT TCTATACAAA CCGCTTAATC ACTATGAGGC 1620
 TTGGCAACAT ATTCAAGTGA CAAATGAGAT TGTACACTC AACTATTGG AACCTGGAC 1680
 AGAATATGAA CTCTGTGTGC AACTGTCTCG TGTGGAGAG GGTGGGGAAG GGCATCTCGG 1740
 ACCTGTGAGA CGCTTCAAAA CAGCTTCTAT CGGACTCCCT CCTCAAAGG GTCTAAATCT 1800
 55 CCTGCCTAAA AGTCAGACCA CTCTAAATTT GACCTGGCAA CCAATATTTC CAAGCTGGGA 1860
 AATAGACTTT TATGTTGAAG TGGAGAGAAG GTCTGTGCAA AAAAGTGATC AGCAGAATAT 1920
 TAAAGTTCCA GGCAACTTGA CTTCGGTGCT ACTTAACAAC TTACATCCCA GGGAGCAGTA 1980
 COTGGTCCGA GCTAGAGTCA ACACCAAGGC CCAGGGGGAA TGGAGTGAAG ATCTCACTGC 2040
 60 TTGGACCTCT AGTGACATT CTCTCTCTCA ACCAGAAAAC ATCAAGATT TCCAACATTAC 2100
 ACACCTCTCG GCTGTGATT CTGAGACAAT ATTGGATGGC TATCTTATT CTCTATTAC 2160
 TATCCGTATC AAGGTTCAAG GCAAGAATGA AGACCAAGC GTTGATGTGA AGATAAAGAA 2220
 TGCCACCATC ATTCAAGTAT AGCTCAAGGG CCTAGAGCCT GAAACAGCAT ACCAGGTGGA 2280
 CATTTTGTGA GAGAACAACA TAGGGTCAAG CAACCCAGCC TTTTCTCATG AACTGTGTAC 2340
 65 CCTCCAGAAA TCTCAAGCAC CAGCGGACCT CGGAGGGGGG AAGATGTGTC TTATAGCCAT 2400
 CCTTGGCTCT GCTGGAATGA CCTGCCTGAC TGTGCTGTG GCTTTCTGA TCATATTGCA 2460
 ATTGAAGAGG GCAATGTGTC AAGGAGAAT GGCCCAAGCC TTCCAAAAGC TGAGGGAAGA 2520
 ACCAGCTGTG CAGTTCAACT CAGGGACTCT GGCCCTAAAC AGGAAGGTCA AAAACAACCC 2580
 AGATCTTACA ATTATCCAG TGCTTGACTG GAATGACATC AAATTTCAAG ATGTGATTGG 2640
 70 GGAGGGCAAT TTTGGCCAAG TTCTTAAGGC GCGCATCAAG AAGGATGGGT TACGGATGGA 2700
 TGCTGCCATC AAAAGAATGA AAGAATATGC CTCCAAGAT GATCACAGGG ACTTTGCCAG 2760
 AGAAGTGGAA GTTCTTTGTA AACTTGGACA CCATCCAAAC ATCATCAATC TCTTAGGAGC 2820
 ATGTGAACAT CGAGGCTACT TGTACCTGGC CATTTAGTAC GCGCCCATG GAAACCTTCT 2880
 GGACTTCTCT CGCAAGAGCC GTGTGCTGGA GACGGACCA GCATTTGCCA TTGCCAATAG 2940
 CACCGCGTCC ACACGTCTCT CCCAGCAGCT CCTTCACTTC GCTGCCGACG TGGCCCGGGG 3000
 75 CATGGACTAC TTGAGCCAAA AACAGTTTAT CCACAGGGAT CTGGCTGCCA GAAACATTTT 3060
 AOTTGGTGAA AACTATGTGG CAAAATAGC AGATTTTGA TGTGCCGAG GTCAGAGGAT 3120

5 GTACGTGAAA AAGACAATGG GAAGGCTCCC AGTGCCTGG ATGSCCATGG AGTCACTGAA 3180
 TTACAGTGTG TACACAACCA ACAGTGATGT ATGGTCTTAT GGTGTGTATC TATGGGAGAT 3240
 TGTAGCTTA GAGGACACAC CCTACTGCGG GATGACTTGT GCAGAACTCT ACGAGAAGCT 3300
 GCGCCAGGCG TACAGACTGG AGAAGCCCTT GAACCTGTAT GATGAGGTGT ATGATCTAAT 3360
 GAGACAATGC TGGCGGGAGA AGCCTTATGA GAGGCCATCA TTTGCCCAGA TATTGGTGTG 3420
 CTTAAACAGA ATGTATGAGG AGCGAAGAC CTACGTGAAT ACCACGCTTT ATGAGAAGTT 3480
 TACTTATGCA GGAATTGACT GTTCTGTGTA AGAAGCGGCC TAGGACAGAA CATCTGTATA 3540
 CCTCTGTGTT CCTTTTCACT GGCATGGGAG ACCCTTGACA ACTGCTGAGA AAACATGCCT 3600
 CTGCCAAGG ATGTGATATA TAAGTGATCA TATGTGCTGG AATTCTAACA AGTCATAGGT 3660
 10 TAATATTAA GACACTGAAA AATCTAAGTG ATATAATCA GATTCTTCTC TCTCATTTTA 3720
 TCCCTCAGCT GTAGCATGCC AGTCCCGTTT CATTTAGTCA TGTGACCAT CTGCTGTGTG 3780
 TTTCCACAGC CTGCAAGTTC AGTCCAGGAT GCTAACATCT AAAAATAGAC TTAATCTCA 3840
 TTGCTTACAA GCCTAAGAAT CTTTAGAGAA GTATACATAA GTTTAGGATA AAATAATGGG 3900
 ATTTTCTTTT CTTTCTCTG GTAATATTGA CTTGTATATT TTAAGAAATA ACAGAAAGCC 3960
 15 TGGGTGACAT TTGGGAGACA TGTGACATTT ATATATTGAA TTAATATCCC TACATGTATT 4020
 GCACATTTGA AAAGTTTGA GTTTTGATGA GTTGTGAGTT TACCTGTAT ACTGTAGGCA 4080
 CCACTTGCAC TGATATATCA TGAATGAATA AATGTCTTGC CTACTCAAAA AAAAAAAA

Seq ID No: 147 Protein sequence:
 Protein Accession #: NP_000450.1

25 1 11 21 31 41 51
 MDLSLASLVLC GVSLLSSTV EGAMDLILIN SLPLVSDAST SLTCTASGMR PHEPITIGRD 60
 FBALNNHQDQ PLEVTDVTR BWAKKVVMKR EKASKINGAY FCBGRVRGEA IRIRTMKMRQ 120
 QASFLPALT/L MTVDKGNVN ISPKKVLIKE EDVVIYKNGS FHSVPRHEV PDILEVHLPH 180
 30 AQPFQAGVYS ARYIGNLFT SAPTRLIVRR CEAGKWSPEC NHLCTACMNN GVCHSDTGBE 240
 ICPFGFMRT CEKACELHTF GRTCKERCSE QEGCKSYVFC LPDPYGCSCA TGWGLQCNE 300
 ACHPGFYGPD CKLRCSCHNG EMCDFQGCCL CSPGWQLQC EREGIPRMTF KIVDLPDHIE 360
 VNSGKFNPIIC KASGWPLFTN EMTLVKPDG TVLHPKDPNH TDHFSVAIPT IHRILPDSG 420
 VVWGSVNTVA GWVEKPFNIS VKVLKPLANA PNVIDTGHNF AVINISSEPY PGDGPISKK 480
 LLYKPVNHYE AWHQIQVTNE IVTLNVLPRR TEYELCVQLV RRGEGGEGHP GFVRRPTTAS 540
 35 IGLPPPPRLN LLPSQTTILN LITWQPIFPSS EDDFYVEVER RSVQKSDQNN IKVFNLTSTV 600
 LIANNLHPRBQ YVVRARVNTK AQGEWSEDLT AWTLEDILPP QPENIKISNT THSSAVISWT 660
 ILDGYSSSTI TIYKVVQGN EDQHVVDVKIK NATIIYQLK GLEPRTAYQV DIFAEINIGS 720
 SNPAFSHELIV TLPSQAPAD LGGKMLLIA ILGSAGMTCL TVLLAPLIL QLRANVQRR 780
 MAQAPQNVRE BPVQFNSGT LALNRKVNKN PDPTIYFVLD WNDIKFQDVI GEGNFGQVLE 840
 40 ARIKKDGLRM DAAIKRMKEY ASKDDHRDFA GELEVLCKLG HHPNIIINLG ACEHRYLYL 900
 AIBYAPHGNL LDFLRKSRVL ETDPAFALAN STASTLSSQQ LHLFAADVAR GMDYLSQKQF 960
 IHRDLARNI LVGENYVAKI ADPGLSRGQE VYVKKTMGR L PVRNMAIESL NYSVYTTNSD 1020
 VWSYVLLWE IVSLGTFYC GMTCAELYEK LPQGYRLEKP LNCDEVIDYL MRQWREKPY 1080
 ERPSFAQILV SLANMLERK TYVNTTLYEK FTYAGIDCSA EBA

Seq ID No: 148 DNA sequence
 Nucleic Acid Accession #: NM_000552.2
 Coding sequence: 311-8752 (underlined sequences correspond to start and stop codons)

50 1 11 21 31 41 51
 AGCTCACAGC TATTGTGGTG GGAAGGGGAG GGTGGTTGGT GGATGTCACA GCTTGGGCTT 60
 TATCTCCCCC AGCAGTGGGG ACTCCACAGC CCTCGGGCTA CATAACAGCA AGACAGTCCG 120
 GAGCTGTAGC AGACCTGATT GAGCCTTTGC AGCAGCTGAG AGCATGGCCT AGGGTGGGCG 180
 55 GCACCATTTG CCAGCAGCTG AGTTTCCAG GAGACCTTGA GATAGCCGCA GCCTCATTTT 240
 GCAGGGGAAG GCACCATTTG CCAGCAGCTG AGTTTCCAG GAGACCTTGA GATAGCCGCA 300
 GCCTCATTTT ATGATTCCCTG CCAGATTTCG CCGGCTGCTG CTGCTCTGG COCTCATTTT 360
 GCCAGGGACC CTTTGTGCAG AAGGAACCTG CCGCAGGTCA TCCACGGCCC GATGAGCCT 420
 60 TTTCCGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG TACAGCTTTG CGGGATACTG 480
 CAGTTACCTC CTGGCAGGGG GCTGCCAGAA ACGCTCCCTC TCGATTATTG GGGACTTCCA 540
 GAATGGCAAG AGAGTGAGCC TCTCCGTGTA TCTTGGGGAA TTTTGTGACA TCCATTGTGT 600
 TGTCAATGGT ACCGTGACAC AGGGGGGACC AAGAGTCTCC ATGCCCTATG CCTCCAAAGG 660
 GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCTT ATGGCTTTGT 720
 GGCCAGGATC GATGCGAGCG GCAACTTTCA AGTCTGCTG TCAGACAGAT ACTTCAACAA 780
 65 GACCTGCGGG CTGTGTGCA ACTTTAACAT CTTTGTGAA GATGACTTTA TGACCAAGA 840
 AGGACCTTGT ACCTCGGACC CTTATGACTT TGCCAACTCA TGGGCTCTGA GCACTGGAGA 900
 ACAGTGTGT GTAGCGGCAT CTCTCCAG CAGCTCATGC AACATCTCT CTGGGGAAAT 960
 CAGAAAGGCG CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT TTGCCCGCTG 1020
 70 CCACCTCTG GTGAGCCCG AGCCTTTTGT GGCCCTGTGT GAGAAGACTT TGTGTGATG 1080
 TGCTGGGGGG CTGGAGTGGC CTTGCCCTGC CCTCTGGAG TAGGCCGGA CCTGTGCCCA 1140
 GAGAGGAATG GTGTGTACG GCTGGACCGA CCACAGCGCG TGCAGCCAG TGTGCCCTGC 1200
 TGGTATCAGG TATAGGCAGT GTGTGTCCCT TTGCGCCAGG ACCTGCCAGA GCCTGCACAT 1260
 CAATGAAATG TGTCAAGAGC GATGCGTGA TGGCTGCAG TGCCTGAGG GACAGCTCTT 1320
 75 GGAATGAAGC CTCTGCGTGG AGAGCACCGA GTGTCCCTGC GTGCATTCC GAAAGCGCTA 1380
 CCTCCCGCG ACCTCCCTCT CTGAGAGCTG CAACACCTGC ATTTGCCGAA ACAGCCAGTG 1440
 GATCTGCAGC AATGAAGAAT GTCCAGGGGA GTGCCTTGT ACTGGTCAAT CCACTTCAA 1500

	GAGCTTTGAC	AACAGATACT	TCACCTTCAG	TGGGATCTGC	CAGTACCTGC	TGGCCCGGGA	1560
	TTGCCAGGAC	CACCTCTTCT	CCATTGTCAAT	TGAGACTGTG	CAGTGTGCTG	ATGACCGGGA	1620
	CGCTGTGTGC	ACCCGCTCCG	TCACCGTCCG	GCTGCTGGC	CTGCACAACA	GCCTTGTGAA	1680
	ACTGAAGCAT	GGGGCAGGAG	TTGCCATGGA	TGGCCAGGAC	ATCCAGCTCC	CCCTCCTGAA	1740
5	AGGTGACCTC	CGCATCCAGC	ATACAGTGAC	GGCCTCGGTG	CGCTCAGCT	ACGGGGAGGA	1800
	CCTGCAGATG	GACTGGGATG	GCCGCGGGAG	GCTGCTGGTG	AAGCTGTCCC	CGTCTACGCG	1860
	CGGGGAAGCC	TGGGCGCTGT	GTGGGAATTA	CAATGGCAAC	CAGGGCGAGC	ACTTCTTTAC	1920
	CCCTCTCTGG	CTGGCAGAGC	CCCGGGTGGG	GGACTTGGGG	AACGCTTGGG	AGCTGCAAGG	1980
	GGACTGCCAG	GACCTGCAGA	AGCAGCACAG	CGATCCCTGC	GCCTCCAAAC	CGCGCATGAC	2040
10	CAGGTCTTCC	GAGGAGGCGT	GCGCGGTCC	GAGCTCCCCC	ACATTGAGG	CGTCCATCG	2100
	TGCCGTGAGC	CGCTGCGCT	ACCTGCGGAA	CTGCGCTAC	GAGCTGTGCT	CCTGCTCGGA	2160
	CGGCGCGAG	TGCTGTGTGC	GCGCGCTGCG	CAGCTATGCC	GCGGCTGCG	CGGGAGAGG	2220
	CGTGCGGCTC	GCGTGGGCGG	AGCCAGGCGG	CTGTGAGCTG	AACTGCGCGA	AAGGCCAGGT	2280
	GTACTGTGAG	TGGGGGAGCC	CCTGCAACCT	GACCTGCGCG	TCTCTCTCTT	ACCGGATGA	2340
15	GGAAATGCAAT	GAGGCGCTGC	TGGAGGCGTG	CTTCTGCCCC	CCAGGGCTCT	ACATGGATGA	2400
	GAGGGGGGAC	TGCGTGCCCA	AGGCCGAGTG	CCCTGTGTAC	TATGAAGGTG	AGATCTTCCA	2460
	GCCAGAGGAC	ATCTTCTCAG	ACCATCACAC	CATGTGCTAC	TGTGAGGATG	GCTTCATGCA	2520
	CTGTACCATG	AGTGAGAGTC	CGGGAAGCTT	GCTGCTGAC	GCTGTCTCTA	CGAGTCCCTC	2580
	GTCTCATCGC	AGCAAAAGGA	GCCTATCCTG	TGGGCCCCCG	ATGCTCAAGC	TGGTGTGTCC	2640
20	CGCTGACAAAC	CTGGCGGCTG	AAGGGCTCGA	GTGTACCAAA	ACGTGCCAGA	ACTATGACCT	2700
	GGAGTGCATG	AGCATGGGCT	GTGTCTCTGG	CTGCTCTGCG	CCCCCGGCA	TGTGCGGCA	2760
	GGAGTGCATG	AGCATGGGCT	GTGTCTCTGG	CTGCTCTGCG	CCCCCGGCA	TGTGCGGCA	2820
	TGAGAAACAGA	TGTGTGGCCC	TGGAAAGGTG	TCCCTGTCTT	CATCAGGGCA	AGGAGTATGC	2880
25	CCCTGTGAGAA	ACAGTGAAGA	TTGGCTGCAA	CATTGTGTGT	TGTGCGGACC	GGAAAGTGAA	2940
	CTGCACAGAC	CATGTGTGTG	ATGCCACGTG	CTCCACGATC	GGCATGGCCC	ACTACCTCAC	3000
	CTTCCAGCGG	CTCAAAATACC	TGTTCCCGGG	GGAGTGCCAG	TACGTTCTGG	TGCAGGATTA	3060
	CTGCGCGAGT	AACCTTGGGA	CCTTTOGGAT	CCTAGTGGGG	AATAAGGGAT	GCAGCCACCC	3120
	CTCATGTGAA	TGCAAGAAAC	GGGTCAACAT	CCTGTGTGAG	GGAGGAGAGA	TTGAGCTGTT	3180
	TGACGGGGAG	GTGAATGTGA	AGAGGCCCAT	GAAGGATGAG	ACTCACTTTG	AGGTGTGTGA	3240
30	GTCTGCGCGG	TACATCATTC	TGCTGTGGGG	CAAAGCCCTC	TCCGTGTGCT	GGGACGSCCA	3300
	CCTGACCATC	TCCGTGTGCT	TGAAGCAGAC	ATACCAGGAG	AAAGTGTGTG	GCCTGTGTGG	3360
	GAATTTGTAT	GGCATCCAGA	ACAATGACCT	CACCAGCAGC	AACCTCCAA	TGGAGGAAGA	3420
	CCCTGTGGAC	TTTGGGAAC	CCTGGAAAGT	GAGCTCCGAG	TGTGCTGACA	CCAGAAAGT	3480
	GCCTCTGGAC	TGATCCCTCT	CCACCTGCCA	TACCAACATC	ATGAAGCAGA	CGATGGTGA	3540
35	TTCTCTCTGT	AGAATCTTGA	CCAGTGACGT	CTTCCAGGAC	TGCAACAGC	TGGTGGACCC	3600
	CGAGCCATAT	CTGATGTCTG	GCAATTTAGA	CACCTGTCTC	TGTGAGTCCA	TTGGGAGACT	3660
	CGCTCTCTTC	TGGACACCA	TTGCTGCTTA	TGCCACGTTG	TGTGCCAGC	ATGGCAAGGT	3720
	GGTGACTGG	AGGACGCCCA	CATTGTGCCC	CCAGAGCTGC	GAGGAGAGGA	ATCTCCGGGA	3780
40	GAAAGGGTAT	GAGGTGAGT	GGGCTATATA	CAGCTGTGCA	CCTGCTGTTC	AAGTCACTG	3840
	TCAGCAACCT	GAGCCACTGG	CCTGCGCTGT	GCAGTGTGTG	GAGGGCTGCC	ATGCCACTCC	3900
	CCCTCCAGGG	AAAATCTTGG	ATGAGCTTTT	GCAGACCTGC	GTTGACCTTG	AAGACTGTCC	3960
	AGTGTGTGAG	TGGGCTGGCC	GGGCTTTTGC	CTCAGAAAG	AAAGTCACT	TGAATCCAG	4020
	TGACCTGTAG	CACCTGCCAGA	TTTGCCACTG	TGATGTGTTC	AACCTCACT	TGGAAGCTG	4080
	CCAGGAGCCG	GGAGGCTTGG	TGGTGCCTCC	CACAGATGCC	CGGTGAGCC	CCACCACTCT	4140
45	GTATGTGGAG	GACATCTCTG	AACCGCGGTT	GCACGATTTT	TACTGCAGCA	GGCTACTGGA	4200
	CTGCTCTCTC	CTGCTGGATG	GCTCTCTCAG	GCTGTCCGAG	GCTGAGTTTG	AAGTCTGAA	4260
	GGCTCTCTCT	GTGACATGTA	TGGAGGGGCT	GCGCATCTCC	CAGAAGTGGG	TCCGCGTGGC	4320
	CGTGTGTGAG	TACCAAGAGG	GCTCCACGCG	CTACATCGGG	CTCAAGGACC	GGAAAGGACC	4380
	GTCAAGAGCTG	CGGGCGATTG	CCAGCCAGGT	GAAGTATGCG	GGCAGCCAGG	TGGCTCTCAC	4440
50	CAGCGAGGTC	TTGAAATACA	CACGTGTTCA	AATCTTCAGC	AAGATCGACC	GGCTGGAAGC	4500
	CTCCCGCATC	GCCCTGTCTC	TGATGGCCAG	CCAGGAGGCC	CAACGGATGT	CCCGGAACCT	4560
	TGCTCGCTAC	GTCCAGGGCC	TGAAGAAGAA	GAAGGTCAAT	GTGATCCCGG	TGGGCATTGG	4620
	GGCCCATGCG	AACCTCAAGC	AGATCCGCTT	CATCGAAGG	CAGGCCCCCT	AGAACAAAGC	4680
	CTTGTGTGCT	AGCAGTGTGG	ATGAGCTGGA	GCAGCAAAGG	GACGAGATCG	TTAGCTACCT	4740
55	CTGTGACCTT	GCCCTGGAAG	CCCTCTCTCC	TACTCTGCCC	CCCCACATGG	CACAAATCAC	4800
	TGTGGGCGCG	GGGCTCTTGG	GGGTTTCCAG	CCTGGGGCCC	AAGAGGAAGT	CCATGGTTCT	4860
	GGATGTGGCG	TTCTGCTCTG	AAGGATCGGA	CAAAATTTGT	GAAGCCGACT	TCAACAGGAG	4920
	CAAGGAGTTC	ATGGAGGAGG	TGATTCAAGG	GATGGATGTG	GGCCAGGACA	GCATCCAGCT	4980
60	CACGCTGCTG	CAGTACTCCT	ACATGGTGAC	CGTGGAGTAC	CCCTTCAGCG	AGGCACAGTC	5040
	CAAAAGGGAG	ATCCTGCAGC	GGGTGGGAGA	GATCCGCTAC	CAGGGCGGCA	ACAGGACCAA	5100
	CATCTGGCTG	GCCCTGGGCT	ACCTCTCTGA	CCACAGCTTC	TTGGTCAGCC	AGGGTGACCG	5160
	GGAGCAGGCG	CCCAACCTGG	TCTACATGGT	CACCGGAAAT	CCTGCTCTCT	ATGAGATCAA	5220
	GAGGCTGCGT	GGAGACATCC	AGGTGGTGCC	CATTGGAGTG	GGCCCTAAAT	CCAAAGTGCA	5280
65	GGAGCTGGAG	AGGATTGGCT	GGCCCAATGC	CCCTATCTCT	ATCCAGGACT	TTGAGACGCT	5340
	GGCTTGGGAG	GCTCTGAGC	TGGTGTGCA	GAGGTGTCTG	TCCGAGAGGG	GGCTGAGAT	5400
	CCCACCCCTC	TCCCTGTGAC	CTGAGTGCAG	CCAGCCCTTG	GAGCTGATCC	TTCTCTCTGA	5460
	TGGCTCTCTC	AGTTTCCAG	CTTCTTATTT	TGATGAAATG	AAGAGTTTCC	CCAAAGCTTT	5520
	CATTTCAAA	GCCATATAG	GGCTCTGCTT	CACCTCAGGT	TCAAGTCTGC	AGTATGGAAG	5580
70	CATCACCACC	ATTGACGTGC	CATGGAAGCT	GGTCCCGGAG	AAAGCCCAAT	TGCTGAGCCT	5640
	TGTGGAGCTC	ATGCAGCGGG	AGGGAGGCCC	CAGCCAAATC	GGGATGCTCT	TGGCTTTTGC	5700
	TGTGGAGTAC	TTGACTTCAG	AAATGCATGG	TGCCAGGCGG	GGAGCTCTCA	AGGCGGTGGT	5760
	CATCTGGTGC	ACGAGCTCTC	CTGTGGATTC	AGTGGATGCA	GCAGCTGATG	CCGCGAGGTC	5820
	CAACAGAGTG	ACAGTGTTC	CTATTGGAAT	TGGAGATGCG	TACAGTGCAG	CCGAGCTACG	5880
75	GATCTTGGCA	GGCCAGCAG	CGGACTCCAA	CGTGGTGAAG	CTCCAGCGAA	TGGAAGACCT	5940
	CCCTACCATG	GTACCTTGG	GCAATTCCTT	CCTCCACAAA	CTGTGCTCTG	GATTGTGTAG	6000
	GATTTCGATG	GATGAGGATG	GGAAATGAGAA	GAGGCCCGGG	GACGCTTGGA	CCTTGCCAGA	

5 CCACTGCCAC ACCGTGACTT GCCAGCCAGA TGGCCAGACC TTGCTGAAGA GTCATCGGGT 6060
 CAACTGTGAC CGGGGGCTGA GGCCCTTCGTG CCCTAACAGC CAGTCCCTCTG TTAAGTGGGA 6120
 AGAGACCTGT GGCTGCGGCT GGACCTGCCCT CTGGGTGTGC ACAGGCAGCT CCACCTGGCA 6180
 CATGTGACCC TTTGATGGGC AGAATTTCAA GCTGACTGSC AGCTGTTCTT ATGTCTTATT 6240
 TCAAAACAAG GAGCAGGACC TGGAGGTGAT TCTCCATAAT GGTGCTTGCA GCCCTGGAGC 6300
 AAGGCAGGGC TGCATGAAAT CCATCGAGGT GAAGCACAGT GCCTCTCCG TCGAGCTGCA 6360
 CAGTGACATG GAGGTGACGG TGAATGGGAG ACTGGTCTCT GTTCTTACG TGGGTGGGAA 6420
 CATGGAAGTC AACGTTTATG GTGCCATCAT GCATGAGGTC AGATTCAATC ACCTTGGTCA 6480
 CATCTTCACA TTCACTCCAC AAAACAATGA GTTCCAACTG CAGCTCAGCC CCAAGACTTT 6540
 10 TGCTTCAAAAG AOSTATGGTC TGTGTGGGAT CTGTGATGAG AACGAGGCA ATGACTTCAT 6600
 GGTGAGGATG GGCACAGTCA CCACAGAGCTG GAAACACCTT GTTCAGSANT GACTGTGCA 6660
 GCGGCCAGGG CAGAGCTGCC AGCCCATCTCT GAGAGGACAG TGTCTTGTCC CGACAGCTC 6720
 CCATCTGAGG GTCTCTCTCT TACCACTGTT TGCTGAATGC CACAAGTGC TGGCTCCAGC 6780
 CACATCTTAT GCCATCTGCC AGCAGGACAG TTGCACCAG GAGCAAGTGT GTGAGGTGAT 6840
 15 CGCTCTTAT GTCCACCTCT GTCGACCAA CGGGGTCTGC GTTGACTGGA GGACACCTGA 6900
 TTTCTGTGCT ATGTCTATGC CACCATCTCT GGTCTACAAC CACTGTGAGC ATGGCTGTCC 6960
 CGGCACCTGT GATGGCAACG TGAGCTCTCT TGGGGACCAT CCTCCGAAG GCTGTTTCTG 7020
 CCCTCCAGAT AAAGTCATGT TGAAGGCGAG CTGTGTCCTT GAAGAGGCTC GCACTCAGTG 7080
 CATGTGTGAG GATGGAGTCC AGCACCAGTT CCTGGAAGCC TGGGTCCCGG ACCACCGACC 7140
 20 CTGTGAGTAT TGACATGCC TCAGCGGGCG GAAGGTCAAC TGCAACAACG AGCCTGCCCC 7200
 CAGGCCCAAG GCTCCACGT GTGGCTGTGT TGAAGTAGCT GCCTCCGCTC AGAATGAGCA 7260
 CCAGTGTCTC CCGAGATATG AGTGTGTGTG TGACCCAGTG AGCTGTGACC TGCCCCAGT 7320
 GCTCACTGT GAACTGTGCC TCCAGCCAC ACTGACCAAC CCTGCGAGT GCAGACCCAA 7380
 25 CTTACCTCTG GCTGTCAGGA AGGAGGAGTG CAARAGAGTG TCCCACTCTT CCTGCCCTCC 7440
 GCACCGTTTG CCCACCTTTC GGAAGACCCA GTGCTGTGAT GAGTATGAGT GTGCTTGCAA 7500
 TGTGTCTAAC TCACAGTGA GCTGTCCCTT TGGGTACTGT GCCTCAACCG CCAACATGA 7560
 CTGTGCTCTG ACCACAACCA CCTGCTCTCC CGACAAGGTT TGTGTCCACC GAAGCACCAT 7620
 CTACCTCTGT GAGCAGTTCT GGGAGGAGGG CTGGATGTG TGACCTGCA CCGACATGGA 7680
 GGATGCCCTG ATGGGCTTCC GGTGTGCCCA GTGCTCCAG AAGCCTCTGT AGGACAGCTG 7740
 30 TCGTGTGGGC TTTCACTTAC TTCTGATGA AGGCGAGTGC TGTGGAAGGT GCCTGCCATC 7800
 TGCTGTGAGT GTGTGACTG GCTCACCGCG GGGGACTCTC CAGTCTTCTT GGAAGAGTGT 7860
 CGGCTCCAG TGCCCTTCCC CGGAGAACCC CTGCTCTATC AATGAGTGTG TCGAGTGA 7920
 GGAGGAGGTC TTTATACAA ACAGGAACGT CTCCTGCCCT CAGCTGAGG TCCCTGTCTG 7980
 35 CCCCCTGGGC TTTCACTGA GCTGTAAAGC CTCAGGCTGC TGCCCAAGCT GTGCTGTGA 8040
 GGCATGGAG GCTGTCATGC TCAATGGCAC TGTCAATTGG CCGGGAAGA CTGTGATGAT 8100
 CGATGTGTGC ACAGCTTCCC GCTGCTATGT GCAGGTGGGG GTCTCTCTGT GATTCAAGCT 8160
 GGAGTGCAGG AAGACCACTT GCAACCCCTG CCCCCTGGGT TACAAGGAAG AAAATAACAC 8220
 AGGTGAATCT TGTGGAGAT GTTGTCTTAC GGTCTGACCC ATTCACTAA GAGGAGGACA 8280
 40 GATCATGACA CTGAAGCGTG ATGAGACGCT CCAGGATGSC TGTGATACTC ACTTCTGCAA 8340
 GGTCAATGAG AGAGGAGAGT ACTTCTGGGA GAAGAGGCTC ACAGGCTGCC CACCTTTGA 8400
 TGAACACAAG TGTCTGGCTG AGGAGGTTAA AATTATGAAA ATTCCAGGCA CCGCTGTGA 8460
 CACATGTGAG GAGCCTGAGT GCAACGACAT CACTGCCAGG CTGCAATATG TCAAGGTGGG 8520
 45 AAGCTGTGAG TCTGAAGTAG AGGTGGATAT CCACTACTGC CAGGGCAAAT GTGCCAGCAA 8580
 AGCCATGTAC TCCATTGACA TCAACGATGT GCAGGACGAG TGCTCTCTGT GCTCTCCGAC 8640
 AGGAGCGAG CCGATGCAAG TGGCCCTGCA CTGCACCAAT GGTCTGTGTT TGTACCATGA 8700
 GGTCTCAAT GCCATGGAGT GCAAAATGCT CCCCAGGAG TGCAAGCAAGT GAGGCTGTCTG 8760
 CAGCTGCTAG GTGTGCTGCT GCTGCTCTGC TTGGCTGTAT GGCCAGGCA GAGTGTCTGCC 8820
 50 AGTCTCTGCT ATGTTCTGCT CTGTGCTGCT TCTGAGCCCA CAATAAAGGC TGAGCTCTTA 8880
 TCTGTCTGCA TGTCTGCTC TTGTGCTCTT CTGAGCCAC AAT

Seq ID No: 149 Protein sequence:
 Protein Accession #: NP_000543.1

55 1 11 21 31 41 51
 MIPARFAGVL LALALILPQT LCAEGTRGRS STARCSLFGS DFVNTFDGSM YSPAGYCSYL 60
 LAGGCGKRSE SIIGDPQNGK RVSLSVYLGE PFDIHLFVNG TVTQGDQVRS MPYASKGLYL 120
 60 ETRAGYKLS GEAYGFVARI DSGSNFPQVLL SDRYFNKTCG LOGNPNIFAE DDFWTQEGTL 180
 TSDPYTFANS WALSSGBQWC ERASPPSSSC NISSGRMQKG LWEQCQLLKS TSVFARCHPL 240
 VDPEFPVALC EKLCEBAGG LECACPALLE YARTCAQEGM VLYGWDHSA CSFVCPAGME 300
 YRQCVSPCAR TCQSLHINEM CQKRCVDCGS CPEGQLLDEG LCVESTTECP VHSKRYPPG 360
 TSLSRDCNTC ICRNSQWICS NEBCPGECLV TGQSHFSPD NRYPTFSGIC QYLLARDCQD 420
 65 HSFSIVTASV QCADDRDAVC TRSVTVRLPG LHSNVLKLGK GAGVAMDDGD IQLPLLLKGL 480
 RIQHTVTFAGV RLSYGEDLQM DWDGRGRLLV KLSFVYAGKT CGLGNYNGN QGDDFLTSPG 540
 LAEPVTVDFG NAMKLHGDQ DLQKQSDPC ALNPRMTFS ERACAVLTSP TFEACHRAVS 600
 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRSVRV ANREPGRCEL NCFKGQVYLQ 660
 CTFPCNLTCT SLSPDEECN EACLEGCFPC PGLYMERGD CVPKAQCPCY YDGEFQPED 720
 70 LRAEGLBCTK TCQNYDLECM SMGCVSGCLC PPGMVRHENR CVALERCPCF HQKEYAPGE 840
 TVKIGCNTCV CRDRKWNCTD HVCDATCSTI GMAHYLTFDG LKYLFPQECQ YVLVQDYCGS 900
 NGPTFRLLVG DVCSSHPSVK CKKRVTLVE GGEIBLFDGE VNVKRPMDK THEFVVEGSR 960
 YLILLKAL SVWVDRHLSI SVVLKQTYQE KVGLCGNFD GIQNNDLTSS NLQVEBDFVD 1020
 75 FQNSWVSSQ CADTRKVPLO SSPATCHNNI MKQTMVDSSC RILTSDFVD CNLVDPEPY 1080
 LDVCIYDTCB CSIGDCACF CDTIAAYAHV CAQHGVVTV RTATLCPQSC BERNLRENGY 1140
 ECEWRYNSCA PACQVTCQHP BPLACFVQCV BGCHACPPG KILDELQTC VDPEDCPVCE 1200

VAGRRFASGK KVTLNPSDFE HCQICHCDVV NLTCEACQEP GGLVVPPTDA PVSPTTLYVE 1260
 DISRPPLDHF YCSRLDLDF LLDGSSRLSE AEFVLKAFV VDMMLRLIS QKWVRVAVVE 1320
 YHDGSHAYTG LKDRKRPSRL RRIASQVYA GSQVASTSEV LKYLTFQIPS KIDRPEASRI 1380
 ALLLMASQEP QRMSRNFVRY VQGLKCKKVI VIPVGIGPHA NLKQIRLIEK QAPENKAFVL 1440
 5 SSVDBLQQR DEIVSYLCLD APEAPPTLP PHMAQVTGVP GLLGVSTLGF KRNSMVLDA 1500
 FVLQGSQKIG EADFNRSKEF MSEVIQRMDF QGDSIHVTVL QYSYMTVEY PFSEAQSKGD 1560
 ILQVRVRIY QGGNRTNGL ALRYLSDHSP LVSQGDREQA PNLVYMTGN PASDEIKRLP 1620
 GDIQVVRIGV GPANVQBLE RIGWPNAPIL IQDFETLPRE APDLVLQRC SGBGLQIPTL 1680
 SPAPDCSQPL DVILLDQSS SFPASYFDEM KSPAKAFISK ANIGFRIATQV SVLQYGSITT 1740
 10 IDVPMNVVPE KAHLLSLVDV MQRREGSPSQI GDALGFAVRY LTSEMHGARP GASKAVVILV 1800
 TDVSVDSVDA AADAARSNRV TVFFIGIGDR YDAALRLILA GPAGDSNVVK LQRIEDLPTM 1860
 TVLGNSEFLHK LCGFVRLCM DEDGNEKRPD DWTLPDQCH TVTCQPDGQT LLKSHRVNCD 1920
 RGLRPSCFNS QSPVKVBETC GCRWTCPCVC TGSSTRHIVT FDGQNPFLTG SCYVLPQNK 1980
 15 EQDLEVLIN GACSPGARQG CMKSEIEVHS ALSVELHSDM ETVNGLRLVS VPYVGNMNV 2040
 NYGAIMHIV RFPHLGHIFT FTPQNNBPQL QLSPKTFASK TYGLGICDE NGANDPMLRD 2100
 GTVTTDWKTL VQWMTVQRPQ QTCQFLEBQ CLVPDSSHCQ VLLFLFAEC HKVLAPATFY 2160
 AICQDSDCHQ BQCVETIASY AHLCRTNGVC VDWRTPDFCA MSCPPLVYN HCEHGCPRHC 2220
 DGNVSSCGDH PSBGCFCPPD KYNLESGCVP EBACTQCIQE DGVQHFLEA WVPDHPQCI 2280
 20 CGLRSGRKYN CTTPCPTAK APTCGLCEVA RLQONADQCC PEYECVDFV SCOLPPVPHC 2340
 BRGLQPTILN PGCRPNFTC ACRKECKRV SPPSCPHRL PTLRKTQCCD EYECACNVN 2400
 STVSCPLVYL ASTATNDGCG TTTTCLPDKV CVHRSTIYVP GQFWBEGCDV CTCTMDSDAV 2460
 MGLRVACQSQ KPCEDSCRSG FTYVLHSGEC CGRCLPSACE VYTGSPRGDS QSSWKSQVGSQ 2520
 WASPENCLIL NECVRVKEEV FIQRNVSCP QLEVFVCPSP QLSCKTSAC CPSCRCRME 2580
 25 ACMNLGTVIG PGKTVMDIVC TTRCMVQVG VISGPKLECR KTCNCPCLG YKSEHNTGEC 2640
 CGRCCLPTACT IQLRGGQINT LKRDDELQDG CDTHFCVNE RGRYFWKRV TGCPFPDEHK 2700
 CLABGGKIMK IPTCCDTCB EPRCNDITAR LQYVKVGSCK SEVEVDIHYC QGKASKAMY 2760
 SIDINDVDQD CSCSPTRTE PMQVALHCTN GSVVYHEVLN AMECKCSPRK CSK

30 Seq ID NO: 150 DNA sequence
 Nucleic Acid Accession #: NM_001508.1
 Coding sequence: 1-1362 (underlined sequences correspond to start and stop codons)

35 1 11 21 31 41 51
 | | | | | |
 ATGCTTCAC CCAGCCTCC GGCAGTGAC TGCTCCCAA TCATGTATCA CAGTCATGC 60
 CCGAGTTTG AGTGGCCAC CTGGATCAAA ATCACCCTTA TTCTGGTGT CCGTATCATC 120
 TTCGTGATGG GCTTCTGGG GAACAGCGTC ACCATTCGGG TCACCCAGGT GCTGCAGAAG 180
 40 AAAGGATACT TGCAGAGGA GGTGACAGAC CACATGGTGA GTTTGGCTTG CTGCGACATC 240
 TCTGTGTTCC TCATGCGCAT GCCCATGGAG TTCTACAGCA TCATCTGAA TCCCTGACC 300
 AGCTCCAGCT ACACCTGTC CTGCAAGCTG CACACTTTC TCTTCGAGGC CTGCAGCTAC 360
 TGTGCTGCTC TGCACTGCT GAGCTCAGC TTTGAGCGCT ACATCGCCAT CTGTACCCCC 420
 45 TTCAGGTACA AGCTGTGTC GGGACCTTGC CAGGTGAAGC TGCTGATGG CTTCCTCTGG 480
 GTACACTCCG CCTGCTGTC ACTGCCCTTG CTGTTTGCCA TGGTACTGA GTACCCCTTG 540
 GTGAACGTGC CCGGCCACG GGGTCTCACT TGCAACCGCT CCAGCACCGG CCAACAGGAG 600
 CAGCCGAGGA CCTCAATAT GTCCATCTGT ACCAACCTCT CCAGCCGCTG GACCGTGTTC 660
 CAGTCCAGCA CTTCGGGCG CTTCGTGGTC TACCTCGTGG TCCTGCTCTC CGTAGCCTTC 720
 50 ATGTGCTGGA ACATGATGCA GGTGCTCATG AAAAGCCAGA AGGGCTCGCT GCGCGGGGCG 780
 ACGCGGCTTC CGCAGTGAG GAAGTCCGAG AGCGAAGAGA GCAGGACCGC CAGGAGGCAG 840
 ACCATCATCT TCTTGAGGCT GATTGTGTGT ACATTTGGCC TATGCTGGAT GCCAACCCAG 900
 ATTCGAGAGG TCATGGCTGC GGCCAAACCC AAGCAGGACT GGACGAGTTC CTACTTCCGG 960
 CGGTACATGA TCTCTCTCCC CTCTCTGGAG ACGTTTCTCT ACCTCAGCTC GGTATCAAC 1020
 CCGCTCTCTG ACACGGTGTG CTGCGAGCAG TTTGGGGGGG TGTCTGTGCA GGTGCTGTGC 1080
 55 TGCCGCTCTG CGCTGACGA CGCCAAACCC GAGAAGCGCC TGCGGTGACA TGCGCACTCC 1140
 ACCACCGACA GCGCCGCTT TGTGCGAGCG CCGTGTCTCT TCGCGTCCCG GCGCCAGTCC 1200
 TCTGCAAGGA GAACCTGAGA GATTTTCTTA AGCACTTTTC AGAGCGAGGC CGAGCCCCAG 1260
 TCTAAGTCCC AGTCATTGAG TCTCGAGTCA CTAGAGCCCA ACTCAGGCGC GAAACCCAGC 1320
 60 AATTCTGCTG CAGAGARTGG TTTTCAGGAG CATGAAGTTT GA

Seq ID No: 151 Protein sequence
 Protein Accession #: NP_001499.1

65 1 11 21 31 41 51
 | | | | | |
 MASPSLEPGSD CSQIIDSHV PEFEVATWIK ITLILVYLII FVMGLLGNV TIRVTQVLQX 60
 KGYLQKEVTD HMVSLACSDI LVFLIGMPME FYSIWNPLT TSSYTLCKL HTFLFEACSY 120
 70 ATLLHLVTLF FERYIAICPE FRYKAVSGPC QVLLIGFVM VTSALVALPL LFMGTBYPL 180
 VNVPSHRGLT CNRSSTRHE QPETSNMISIC TNLSSRWTFP QSSIFGAPVV YLVLLSVAF 240
 CMWNMQVLM KSQKSLAGG TRPPQLRKSE SEESRTARRQ TIIFRLRIVV TLAVCNMPNQ 300
 IRRIMAAAPK KHDWTRSYFR AYMILLPFSE TFFYLSVIN PLLYTVSSQ FRRVFPVQLC 360
 75 CRLSLQHAKH EKRLRVHAHS TTDSARFVQR PLLFASRRQS SARRTEKIFL STQSEAEFP 420
 SKSQSLLES LEFNSGAKPA NSAAENGPFQE HEV

Seq ID NO: 152 DNA sequence

Nucleic Acid Accession #: none found

Coding sequence: 3-65 (underlined sequences correspond to start and stop codons)

```
5 1 11 21 31 41 51
   | | | | |
   | | | | |
TATTATTT GTGTAACCTA TATTCTGCTT ATAGAGAGTC TCTGAGACTA AAATTGACAA 60
CTTGAAGAGT ATTCCAAGGA ATATTATGAA AATAGGGCAA CATGGACTGT TTAAGATCTC 120
CATGTAAATG AAATTCATGC AAGGAAACAA CTCATAGAAA AGATAAATAT GGATGCCCTT 180
10 CACATCTTAT CAACCTCGTA ACTTTTGGTG CTTGCTGAAT CAGTCCATGA AAAGCTACAG 240
CCGCTCTTT GGGAAATGCTA CATACCCATT TCTGGTATTT AAAAAATATC TAGGAGGAGC 300
TAAATGACAA AACACAGCAG TGTTTTGAGG GAGAAAGGAC CATCATTTAT AATGCTCTGT 360
ACATATACC AGAGCTGCTT GGAATAATTAA AGGCCACTTG TGGCTTTTTC CTACCAACTG 420
ATACGTTTAA ATTTGCCCTA GGATTSAGCT AACAGCAAAA AAAAAAAAAA AAAAAAAAAA 480
15 GAGAGAAAGA AAGGAGKAAA CAGTGGTAAT AAAAAATCC ATCTGCTTTC TTGCTATGTT 540
AATATTAATA AATCATAATA TGACAAGACC CTCACTGAAT AAGAGTATTT TCAGTCATCA 600
GAGCCAGCT GTTGGTAGGC ATTAATGAGT TTAATAATTGT TCTCAATTGA AAAAAATCA 660
CAGTATTTTG CCAAAACCAA AGTAATATA ATACTGTGTC CTCTGTAAAT TTTTGAGAA 720
GTGTTTATAA AGGCGATATT TACATAAATT CTACTTTATT CCTCAACTTC TTTGATGAAT 780
20 GTAACCCAT TTTACTTCTT TAAAAAGTCT CAATTCAGC TGGATTAGCC AGCTCAGCAT 840
AATCACTAG ACAGTGGTIT GTTAAATTTA GCAGCATACT TCGTCCCAT TCTAATTAAA 900
GTGATGAGTT CTTGAATCCC AGAGAAATAA TGCTTAGGAA CTCTCTCAAC TCTGCTTGGC 960
TTGGCTTAGA GAAGTGGCCA TTTTATCAAC AGGAAAAAAA AAAATTTTCT CTACTACAC 1020
CCGTTTGCTT TCTGAAAAAC AGCAAGTTAT TTCTTTATAT AATTATCATT TTATTATTT 1080
25 ATGGAATATT AATTATTATA TTAATAGCCT ATTATGTGTT CTCACTTGCT TCTCTAAGTA 1140
ATATTTTGAAG ATAAATGTGTT GAATAAAACC ATGGATTATA GAGAAAAGTC AAAATATATG 1200
TGTAATATTT AATTATTTTA TAAGTTTAT AATAAAGTAT TCCATTCTCT TATCTT
```

30 Seq ID No: 153 Protein sequence:

Protein Accession #: none found

```
35 1 11 21 31 41 51
   | | | | |
IILCKLYSAY RESLRKLKLT
```

Seq ID NO: 154 DNA sequence

Nucleic Acid Accession #: none found

40 Coding sequence: 1-36 (underlined sequences correspond to start and stop codons)

```
45 1 11 21 31 41 51
   | | | | |
CTGGATGATA TGAAGAAGAT GGATGGGTTA AGGTAAAGG CTGATCAGAG ATGGGTTCTCT 60
CTCAAGGTTA AATATAGTTA AGTGCCAGAA GAAAAGGTGG GCACCAGCGA ATTAAGAACC 120
ATCTTTGAAAT GGTCCCTTGG GTTAAATACT TAACTTTGT CATCAGTGTG TGCATTATG 180
AAATGAAGAG GAATTCACCTA ATATGCTAGC TGATCTTTTG TTTGTCATGA AAAGAGTTAC 240
TGTTGTGTAG TTCTCTGTTT CAGGGCTGCC TTTGCTCCAC AAAGCACTGA GAAGCAGTGG 300
50 CCTGTACAA CCATCTGACC TCTCAACACT GTGTAAAGG CTAAACCCGC CCAGCGAACC 360
TTCTCTGGAG ATATAAAATA CATAGGTTTA GGCTGGCAAA AAAAAAAAAA AAA
```

Seq ID NO: 155 Protein sequence:

Protein Accession #: none found

```
55 1 11 21 31 41 51
   | | | | |
LDDNEEMDGL R
```

Seq ID NO: 156 DNA sequence

Nucleic Acid Accession #: NM_032961.1

65 Coding sequence: 827-3949 (underlined sequences correspond to start and stop codons)

```
70 1 11 21 31 41 51
   | | | | |
CAGGCTCAGA GGCTGAAGCA GGAGGAAGGA AGGACTGGAA GGAAGAAAGAG ACAGGTTAGA 60
GGGAAAGAGG CTGGGAAGA AAACAGCAGA AAAGAAACTG CTCATTACAC TTACAGAGAG 120
GCAAGTAACG GTGGAGATGA GGACAGAGGG AACCAAGACT CTGAAGACA AAAAATACAA 180
ATAGAGCGAA AGAGGAAAAA AATGTCAGAA AGAACATCCA TCCGGAGAAA TGAAGAGAAT 240
GAAAGTTTAA AACTGCAGAG CCGTCTGTG CTTTTCGGC ACAAAATTAT ATCGCTGATT 300
75 TTAAGCCCTT TTGCATTGCG CAGCCGTGTA CATTAGAGG CATOTTTAAC GGTGCCAACA 360
GCATCTCCTT TTCTTCTGCC TCTTCTCTT CTCTCTCTC CTCTCTCTC TCTCTTTTT 420
```

CCTCCTCCTC GTTCTCCTCC CATCAGCAAG AAGACAAACC GAGGACAGTC TTGAATATATC 480
 GAAATTTCTT CTTTGGGATT TGCCAGCGCC AAGACTGTGG GAATAAAGGA CGCTGACTAT 540
 TGTATTATTG TTATTTTATT AATTAGTCAG TGGAAAGATT ACAGATGAGG AAAGGGGACG 600
 CCTCTCACCC TTCTGTGCT AAGATTAA AAAAAATGAG GCTGGATTGC GGGAGCTCT 660
 5 AATATGAAGC AAAAGGAGTA AGATTTTAA AGACAGAAAG CCACAGGAGC CCCCACGTAG 720
 CGCACTTTTA TTTGTATTTT TTCAGATTTT TTTTGTGTTT GTGGTGGTGG GGGAGGTGAT 780
 TGGGTGGCTG ACTGGCTGGG GGAAGCTACT TCCTTTCTCT TTGGAGATGA TTGTCTATT 840
 ATGTGTTGCC TTGCTCTGGA TGGTGGAAAG AGTCTTTTCC CAGCTTCACT ACACGTACA 900
 GAGGAGCAGC GAACATGGCA CTTTCTGGG GAATATGCT GAAGATCTGG GTCTGGACAT 960
 10 TACAAAACCT TCGGCTCGCG GGTTCAGAG GGTGCCCAAC TCAAGGACCC CTACTTAGA 1020
 CCTCAACCTG GAGACAGGGG TGCTGTACGT GAACGAGAAA ATAGACCGCG AACAAATCTG 1080
 CAACAGAGCG CCTCCTGTG TCCTGCACCT GGAGGTCTTT CTGGAGAACG CCTGGAGCT 1140
 GTTCCAGGTG GAGATCGAGG TGCTGGACAT TAATGACAA CCCCCCTCTT TCCCGAGGCC 1200
 AGACCTGAGC GTGGAATCT CTGAGAGCGC CAAGCCAGGC ACTCGTTCC CCTTGGAGAG 1260
 15 CGCATTCGAG CCAGACGTGG GCACCACTC CTTGCGGAC TACGAGATCA CCCCACAG 1320
 CTACTTCTCC CTGAGCGTGC AGACCAAGGG GGATGGCAAC CGATTGCTG AGCTGGTGCT 1380
 GGAGAGCCCA CTGAGCCGAG AGCAGCAAGC GTGACCGC TACGTGCTGA CCGCGTGA 1440
 CGAGGAGGAT GGGGAGGAG TAGGAGAAAG AGGGGAGGT GGGGGGAGG CAGGCGTCC 1500
 CCCCCAGCTC CAGCGCACCG GCACGCGCT ACTCACATC CAGTGTCTG ACTCCAATGA 1560
 20 CAATGTGCGC GCTTTGAGC AACCGTCTA CACTGTGTCC CTACAGAGA ACTTCCCC 1620
 AGGCACTCTC GTGATCCAGC TCAAGCCAC CGACCGGAC GAGGCGCAGA ACGGTGAGT 1680
 CGTGTACTCC TTCAGCAGCC ACATTTGCC CCGGCGCGG GAGCTTTTCC GACTCTGCC 1740
 GCGCACTGGC AGACTGGAGG TAAGCGGCGA GTTGGACTAT GAGGAGAGCC CAGTGTACCA 1800
 AGTGTACGTG CAAGCCAGG ACCTGGGCCC CAACGCGTG CAGCGCACT CCAAGGTGCT 1860
 25 AGTCCAGATA CTGGATGCTA ATGACAACGC GCCAGAGATC AGCTTCAGCA CCGTGAAGGA 1920
 AGCGGTGAGT GAGGGGCGCG CGCCGCGCAC TGTGGTGCC CTTTTCAGCG TGACTGACG 1980
 CGACTCAGAG GAGAAATGGG AGGTGCACTG CGAGCTACTG GAGAGCGTGC CTTTCCGCT 2040
 CAGTCTTCTC TTTAAGAAAT ACTACACCAT CGTTACCGAA GCGCCCTG GACGAGAGGC 2100
 GGGGAGCTCC TACACCTGA CTGTAGTGGC TGGGAGCGG GCGGAGCGCT CGCTCTCCAC 2160
 30 CAGTAAGTGC ATCCAGGTAC AAGTGTGCGA TGTGAACGAC AACGCGCGC GTTTCAGCCA 2220
 GCGCGTCTAC GAGCTGTATG TGACTGAAAA CAAGCTGCCT GCGGCTTACA TCTACGGGT 2280
 GAGCGCCACC GACCGGATG AGGCGGCGAA CGCCAGCTT GCTACTCTA TCTCGAGT 2340
 CCAGATCCAG GGCATGAGCG TCTTCACCTA CGTTCTATC AACTCTGAGA ACGGCTACT 2400
 35 GTACGCGCTG CGCTCCTTCG ACTATGAGCA GCTGAAGGAC TTCAGTTTTC AGGTGGAAGC 2460
 CCGGAGCGCT GCGAGCGCCC AGGCGCTGCG TGTAAACGCC ACTGTCAACA TCTCATAGT 2520
 GGATCAAAAT GACAAACGCC CTGCCATCGT GCGCGCTCTA CCAGGGCGCA ACGGAGCTCC 2580
 AGCGGTGAGG GTGCTGCCCC GCTGCGGGA GCGGGTTAC CTGTCAACCC GCGTGGCGC 2640
 CGTGGACGCG GACGACGCGG AGAACGCGCG GCTCACTTAC AGCATCGTGC GTGGCAACGA 2700
 40 AATGAACCTC TTTGCGATGG ACTGGCGCAC GCGGGAGCTG CGCAGAGCAC GCGAGTCCC 2760
 GGCAGAGCGC GACCCCGAGC GGCCTTATGA GCTGTGATC GAGGTGCGCG ACCATGGGCA 2820
 CAGTCTTCTC TCTCCACCG CCACCTGTGT GGTTCAGCTG GTGGATGGCG CCGTGGAGCC 2880
 CCAGGGCGGG GCGCGGAGCG GAGGCGGAGG GTCAGGAGAG CACCAGCGCC CCACTCGCTC 2940
 TGGCGGCGGG GAAACCTCGC TAGACCTCAC CTCTACCTC ATCATCGCGT TGGGCTCGGT 3000
 45 GTCTCTATC TTCTGTCTGG CCATGATCGT GCTGGCGGTG CGTGGCCAA AAGAGAGAA 3060
 GCTCAACATC TATACTTGTG TGCCAGCGA TTGCTGCTC TGCTGCTGCT GCTGCGTGG 3120
 CGGAGGTTCG ACCTGCTGTG GCGGCCAAGC CCGGGCGCGC AAGAAGAAAC TCAGCAAGTC 3180
 AGACATCATG CTGGTGAGCA GCTCCAATGT ACCCAGTAC CCGGCCCAGG TGCCGATAGA 3240
 GAGGTGCGGG GGCTTTGGCT CCCACCAACA CAACAGAAAT TACTGTATC AGGTATGCT 3300
 50 GACCCCTGAG TCCGCCAAGA CCGACCTGAT GTTCTTAA GCTTCAGGCC CTTCGCGGAG 3360
 TACGGACACT GAGCACAAAC CCTGCGGGC CATCGTACC GGTACACCG ACCAGCAGCC 3420
 TGATATCATC TCCAACGAA GCATTTTGTG CAACGAGACT AAACACACG GAGCAGAGCT 3480
 CAGCTATCTA GTTGACAGAC CTCGCCAGT TAACAGTTCT GCATTCCAGG AAGCCGACAT 3540
 AGTAAGCTCT AAGGACAGTG GTATGAGAGA CAGTGAACAG GGAGATAGTG ATCATGATG 3600
 55 CACCACCGT GCCCAGTCAG CTGGTATGGA TCTCTTCTC AATTGCACTG AGGAATGTAA 3660
 AGCTCTGGGC CACTCAGATC GGTGTGAGT GCCTTCTTT GTCCCTTCTG ATGAGCGCA 3720
 GCTGTCTGAT TATCGCAGCA ATCTGCACTG TCCTGGCATG GACTCTGTT CAGACACTGA 3780
 GGTGTTTGA AACTCCAGAA CCCAGCTGG GCGAGAGCG TCCTTTTCCA CTTTGGCAA 3840
 AGAGAAGGCC CTTACAGCA CTCGGGAGG GAAGGAGCTG GATGGACTGC TGAATAATC 3900
 GCGAGCGCCT TACAAACCA CATATTGAC ACGGAAAAGG ATATGCTAGT CAATTCTACA 3960
 60 GGACTTACCT GAAGCAGCAT GATTGACACA AAGTCGACCA ACAGAGCAT CAATTTTCA 4020
 ACTTCATAT CTTGCCATC CAGTTAGTCA TGTGTAACTG AGTATTAGAT TTGGATGGA 4080
 GTCATCATGG CCAATTATAG GACCTAATG CTCTCAGCAG GCCTGAGAAA TGAATTGAA 4140
 TGTGAGAAC TGTAGAACT TTAGAGGCAA CAGATTTGTC CTCCCGATC AGTGTGTGCC 4200
 TTTTACAGC ACTATCTATC TTTCTCTCT CAATGTGAC TGAGCCCTTT AGATGTTAT 4260
 65 TTTTCAACG AGAAGCCAGT CATAAGATA AAGGAAATTT GTGCATTATA AATGCAATAT 4320
 CACTGTTTTA AACTGACTG TTTTATATTA TTTTGTGTG ATCAAGTGT CCGCAAGCTA 4380
 TTTCAACTTT ACAGAGAAA TTGTGATTAT GTTCTTTCA CTTGTGGGT ATAAAAATG 4440
 TTGTATTCTG AAGACCCACA AAATATCAAA GACATTTCT AGTTATATA CCGTGTGCA 4500
 AAGTGTTTAC TGTACTATTT CAAAGCTTCT AAATAAATAT AAATATATA TATTATATTA 4560
 70 TATAATTTTC CTAATAATGT GTACAACTCA GTTGGTTTTT AAATGGATGC ATACAGTCCA 4620
 CATCATACAA TAAAAATAAA GGTAATTGAG GTTCCCAAG ACAGAACTTAC TAAGAAAAA 4680
 TCACTAATAG TTTTCTCCCA ATTTCATAT CTACTCAAC CGTGTTTTTT CTTGTTTAA 4740
 AGAATAATGAT GCTCTAAGCT ACAGAAATTT GTCAAAACT CATATTGAAT TTTCAATGCC 4800
 75 AAGAGTGTAG TATTGATGT TATCAGACAG AGCACTGACT ATGTACTATC AAATATCTA 4860
 ACAATCTGCA TAAGTCTGAT TCTATTCTA TGACTTTGAA TTTAGAAATCA CTTAAAGCTT 4920
 TTATAAAGAA TCGATAAATT CACCTGTATT TGTGTTAGA AAAAAACTGG GTGCTGTGAC 4980

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75

AATTTGTGTT GTAAATATG TAATTGAAGA TTAATATTTT AGAAGTCAT CAGTCATATC 5040
 ACTCACACAG AATTTTATTT TACATAGTTT TGTGACTTAA TTACACATGA ATATAAAATC 5100
 TATAATCTTA TATGAATATA TAGAGATATA GAAACATCTG AACTGTGAAA GAATAACTAT 5160
 AAAAATATGAA AGCTCTAAAT TTAATAATAA TTTAGAGATA GAATCATGGT ACATATTTGT 5220
 TTCAATATTC CATGTAAAAA TTTTATAGCT TAAATGTAGT CAGTGTGTTGA TTAATGAAAA 5280
 AATCTTTCAT GAGTCAGCCT TCAAAAGTTA AGCTTGCCCT TTAATTTTAT GTCAACAATA 5340
 TTAATATTA AATTTAGTAA GACGCAAAAA AAAAAAATAA AAAA

Seq ID No: 157 Protein sequence
 Protein Accession #: NP_116586.1

1 11 21 31 41 51
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85
 90
 95
 100

MIVLLLPALL NMVEGVFSQL HYTVQEEQEH GTFVGNIAED LGLDITKLSA ROPQTVPNSR 60
 TPYLDLLEST GVLVYNEKID REQICKQSPS CVLHLEVPLE NFLELPQVRI EVLDINDNPF 120
 SPFEPDLTVB IESBAPQTR FPLESAFDPD VGTNSLRDYE ITPNSYFSLD VQTQDGNRF 180
 AELVLKPLID RQQQAVHRYV LTAVDGGGGG GVGEGGGGGG GAGLEPQQOR TGTALLTIRV 240
 LDSDNVPAP DQPVYVSLP ENSPPGTLVI QLNATDFDEG QNGEVVYSFS SHISPRAREL 300
 PGLSPRTGRL FVSGELDYEE SPVYQVYVQA KDLGPHAVPA HCKVLVRVLD ANDNAPEISF 360
 STVKEAVSGB AAPGTVVVALF SVTDRDSEEH QVQCELLD VPRLLKSSFK NYTIVTRAP 420
 LDREAGDSYT LTVVARDGE PALSTSKSIQ VQVSDVNDNA PRFSQPVYDV YVTRNVPGA 480
 YIYAVSATDR DBGANAQLAY SILECQIQGM SVPTVYSINS ENGLYALRS FUYQLKDFS 540
 PQVBARADGS PQALAGNATV NILIVDQNDN APAIVAPLPG RMTGPAKRVL PRSAPGYLL 600
 TRVRAVDADD GBNRLTYSI VRGEMNLFR MDWRTGELRT ARRVPAKRDV QRPYELVIEV 660
 RDHQPFLLS TATLVVQLVD GAVEPQGGGG SGGGSGGEHQ RPSRSGGET SLDLTLILII 720
 ALGVSFIFL LAMIVLAVRC QKEKCLNIYT CLASDCLCC CCCGGGGSTC CGRQARARKK 780
 KLSRSDIMLV QSSNVPSNPA QVPIESGGF GSHHHNQNYC YQVCLTPESA KTDLMFLKPC 840
 SPSSRSDTSH NPCGAIVTGY TDQDPDIISM GSILSNETH QRAELSYLD RPRRVNSSAF 900
 QEADIVSSKD SGHGDSEQGD SDHDATNRAQ SAGMDLPSNC TEECKALGHS DRCMPSFVP 960
 SDGRQAADYR SNLHVPQMDS VPDTEVFETP BAQPGAERSF STFGKEKALH STLRKKELOG 1020
 LUTNTRAFYK PPYLTRKRIC

Seq ID NO: 158 DNA sequence
 Nucleic Acid Accession #: NM_022159.1
 Coding sequence: 70-1890 (underlined sequences correspond to start and stop codons)

40
 45
 50
 55
 60
 65
 70
 75

1 11 21 31 41 51
 GTGAATTTA AACTCCAGTC CTGTGGCGAA AATGCTAATT GCACTAACAC AGAAGGAAGT 60
 TATATTGTA TGTGTGTACC TGGCTTCAGA TCCAGCAGTA ACCAAGACAG GTTTATCACT 120
 AATGATGGA CCGTCTGTAT AGAAATGTG AATGCAAACT GCCATTAGA TAATGTCTGT 180
 ATAGCTGCAA ATATTAATAA AACTTTAACA AAAATCAGAT CCATAAAGA ACCTGTGGCT 240
 TTGCTCAAG AAGTCTATAG AAATCTGTG ACAGATCTTT CACCAACAGA TATAATTACA 300
 TATATAGAAA TATTAGCTGA ATCATCTTCA TTAATAGGTT ACAAGAACAA CACTATCTCA 360
 GCCAAGGACA CCGTTTCTAA CTCAACTCTT ACTGAATTTG TAAAAACCGT GAATAATTTT 420
 GTTCAAGGG ATACATTGT AGTTTGGGAC AAGTTATCTG TGAATCATAG GAGAACACAT 480
 CTTACAAAAC TCATGCACAC TGTGTGAACA GCTACTTTAA GGATATCCCA GAGCTTCCAA 540
 AAGACACAG AGTTTGATAC AAATCAACG GATATAGCTC TCAAGTTTTT CTTTTTGAT 600
 TCATATAACA TGAAACATAT TCATCCTCAT ATGAATATGG ATGAGAGCTA CATAAATATA 660
 TTTCCAAAGA GAAAGCTGC ATATGATTCA AATGGCAATG TTGCAGTTGC ATTTTATAT 720
 TATAGAGATA TTGTCTCTTT GCTTTCATCA TCTGACAACT TCTTATTGAA ACCTCAAAAT 780
 TATGATATT CTGAAGAGGA GAAAGAGTC ATATCTTCAG TAATTTCACT CTCAATGAGC 840
 TCNAACCCAC CCACATATA TGAACCTGAA AAAATAACAT TTACATTAA TCAATGAAAG 900
 GTCACAGATA GGTATAGGAG TCTATGTGCA TTTTGAATT ACTCACCTGA TACCATGAAT 960
 GGCAGCTGGT CTTAGAGGG CTGAGAGCTG ACATATCTCA ATGAGACCCA CACCTCATGC 1020
 CGCTGTAATC ACCTGACACA TTTTGCAATT TTGATGTCTT CTGGTCTTTC CATTGGTATT 1080
 AAAGATTATA ATATTCTTAC AAGGATCACT CAACTAGGAA TAATTATTTC ACTGATTGT 1140
 CTTGCCATAT GCATTTTAC CTTCTGGTTC TTCAGTGAAA TTCAAGCAC CAGGACAACA 1200
 ATTCACAAA ATCTTTGCTG TAGCCTATTT CTTGCTGAAC TTGTTTTTCT TGTGGGATC 1260
 AATACAAAAT CTAAATAGCT CTTCTGTTC ATCATGCGG GACTGTACA CTACTTCTTT 1320
 TTAGCTGCTT TTGCATGGAT GTGCATTGAA GGCATACATC TCTATCTCAT TGTGTGGGT 1380
 GTCATCTACA ACAAGGGATT TTTGCACAAG AATTTTATA TCTTTGGCTA TCTAAGCCCA 1440
 GCCGTGGTAC TTGATTTTC GGCAGCACTA GGATACAGAT ATTATGGCAC AACCAGGTA 1500
 TGTGGCTTGA GCACCGAAAA CAACTTTATT TGGAGTTTAA TAGGACCAGC ATGCCAATC 1560
 ATTCTGTGTA ATCTCTTGGC TTTTGGAGTC ATCATATACA AAGTTTTTTC TCACACTGCA 1620
 GGGTGAAGAC CAGAAGTTAG TTGCTTTGAG ACATAAGGT CTTGTGCAAG AGGAGCCCTC 1680
 GCTCTCTGT TCCTTCTCGG CACCACCTGG ATCTTTGGGG TTCTCCATGT TGTGCACGCA 1740
 TCAGTGGTTA CAGCTTACCT CTTACAGTGC AGCAATGCTT TCCAGGGGAT GTTCATTTT 1800
 TTATCTCTGT GTGTTTTATC TAGAAAGATT CAAGAAGAAT ATTACAGATT GTTCAAAAAT 1860
 GCCCTCTGTT GTTTTGGATG TTTAAGGTAA ACATAGAGAA TGGTGGATAA TTACAACTGC 1920
 ACAAAATAAA AATTTCCAAG CTGTGGATGA CCAATGTATA AAAATGACTC ATCAAATAT 1980
 CCAATATTA TACTACTAGC AAAAAATATT TTAATCAGT TTTCTGTTT ATGCTATAGG 2040
 AACTGTAGAT AATAAGTAA AATTATGTAT CATATAGATA TACTATGTTT TTCTATGTGA 2100
 AATAGTCTGT TCAAAAATAG TATTGCAGAT ATTTGGAAG TAATTGTTT CTCAGGAGTG 2160

5 ATATCACTGC ACCCAAGGAA AGATTTTCTT TCTAACACGA GAAGTATATG AATGCTCTGA 2220
 AGGAAACCAC TGGCTTGATA TTCTGTGAC TCGTGTGCC TTGAAACTA GTCCCTACC 2280
 ACCTCGGTAA TGAGCTCCAT TACAGAAAGT GGAACATAAG AGAATGAAGG GGCAGAAATAT 2340
 CAAACAGTGA AAAGGGAATG ATAAGATGTA TTTTGAATGA ACTGTTTTTT CTGTAGACTA 2400
 GCTGAGAAAT TGTGACATA AAATAAGAA TGAAGAAAC ACATTTTACC ATTTTGTGAA 2460
 TTGTTCTGAA CTAAATGTC CACTAAACA ACTTAGACTT CTGTTTGCTA AATCTGTTTC 2520
 TTTTCTAAT ATTCTAAA

10 Seq ID No: 159 Protein sequence
 Protein Accession #: NP_071442.1

15 1 11 21 31 41 51
 MCVPGFRSS NQDRFITNDG TVCIENVNAN CHLDNVCLAA NINKTLTKIR SIKEPVALLQ 60
 EVYRNVVDL SPTDIIITYE ILABSSSLG YKNTISAKD TLSNSTLTFE VKTVNPFVQR 120
 DTFVVDKLS VNRHRTHTK LMTVEQATL RISQSPQKT BFDNSTDLA LKVFPFDSYN 180
 MKHHPHNM DGDYINIPFK RKAAYDSNGN VAVAFLYKS TGPLSSSDN FLKPKQNYDN 240
 SEEBEVRIS VLSVMSNNP PTLVLEKIT PTLSHKVTD RYRSLCAFWN YSFDTMNGSW 300
 20 SSGECSLTYS NETHS6RCN HLTHFAILMS SGPSIGIKDY NILTRITQLG IISLILCLAI 360
 CIPTFWPFS IQSKRTTIHK NLCCSLFLAR LVFLVGINTN TKMLFCSILA GLLHYFFLAA 420
 FAWMCEBGIH LYLIVGVIVY NKQFLHKNFY IFGYLSPAVV VGFSALGYR YGTTKVCWL 480
 STENNFIMSF IGPACLILV NLLAPGVIIY KVPRTAGLK PEVSCFENIR SCARGALALL 540
 25 FLGTTWIFG VLVVHVASV TAYLPTVSN PQGMFIFLFL CVLSRKIQEE YYRLFKNVPC 600
 CFGCLR

30 Seq ID NO: 160 DNA sequence
 Nucleic Acid Accession #: none found
 Coding sequence: 1-216 (underlined sequences correspond to start and stop codons)

35 1 11 21 31 41 51
 TGTCTGCTTA TGCGGTGGCT CGCTGCTCAG AACAGSATGG CAGAGATGAG CACCACCATC 60
 AAAAACTCAA GGACCACTGC TGTGGGTCCA GTCATCTGTT TCATGGAATT CACCAGTCTG 120
 GTATCTTCAA AATCCAGAGG GATGATGGCA GATGGCAGGA AGGAGGAAGA GGGTAATCTG 180
 GAGAGATTTC CTGACCTACT CTGCTGCTGT GATTAACAA CCACAGGAA ATTTTATGTA 240
 40 CACTGTTCTC CTGAGCTCCT CCCTTTCCTC GGGGAAGAAA AGCATTGAAA CTCACAAAAT 300
 AAAGTGTAT TGGCTGGAG TGAGGTCTCA TGCTGCTTA TGCGGTGGCT CGCTGCTCAG 360
 AACAGGGAAC CATGGAGAT ACTATTACT CTTTGAAGGC TTACAGTGGG ATGAATTCAA 420
 ATACGACTTA TTGAGGAAT TGAAGTTGAC TTTATGAGGC TGAAGAAT CTCTTGGAG 480
 AAAAAAGAC TGGTACTTCT GAATTAACCA AATCAGAT ATTCTGAAGA TGATCTTACA 540
 45 AAGCTGTCTG TTCTACAAA GGCTGCTGAT GATTCTACA AAGCTGTCTG TAGTGTGCT 600
 GTGCGCTCTG CTAAAAAG TAGAAAAAC ATTGATGCGC CATGTTCAAC CCAACCTCCC 660
 TGCCTAAAGG CTCAGGACC ATCTTGAAG AGGAAGGCGC GTGAGATTGT AAGAGCCGAA 720
 TTAGGGGGAT GGAGTGTGGA GAATAAGGAC ACTTCATCTT GGATGCTCAC CTGCCAAATT 780
 GACTTCTGAT GAAAGCCAGC TCCAGAAATG TGCCTACAGT TACTACTTTC ACCTAAACCC 840
 TGCCCTTAGT CAAATCCTTC TCTTCTCTA AGCAATCAAC TTCAATCTCT TGTATAACCC 900
 50 ACAGTATAAA AGGGCTTTA TACCATCTA TCCTATTGCA TGTAAAGCCTT GGGTCTGGGA 960
 GGTAAAGTGT TGGGATTTCA CCATCTCATC TCCTGCCAC CCAACATGC CTGCTCTCT 1020
 TTAAGCAATA TTAATGTTT GTACTTCA

55 Seq ID No: 161 Protein sequence
 Protein Accession #: none found

60 1 11 21 31 41 51
 CLIMRWLAQ NRMAEMSTTI KNSRTSAVGP VICFMEFTSL VSSKSRMMMA DGRKEEENL 60
 EEFDDLCCD D

65 Seq ID NO: 162 DNA sequence
 Nucleic Acid Accession #: none found
 Coding sequence: 1-159 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 GAGACCTCC AGAGGCAGGG CCCAGGATTG AAGAGGGAAG CCTGCTCCA CACGTGTTC 60
 TCAGGAAGGA CCCACAGACT GCTGCTCCTG GAGGCTCTC GGTATTGGA TGTGTGTTG 120
 TTCCATAAAC CCTCAGAGGG TCACCTGGAG ACCCGCTAAA ATGCAGGTTT TTGGGCCACA 180
 CCTAGACCT TCTACCGAC CCAGGGAGTG GGGCCAGGA AGCTGCATTG GACAGATATC 240
 CCCGTGTGAT CATCATGCAC ACAGGAGTGA GAGAACCAGT GTTCTCCCGG GGCAGAAAGG 300
 75 AAGCTCGTGT GCAGGACACC TCACCTCTCC TTTCCCATTC CCTGCCAGG CTCTCCCTGC 360
 TGACATTGTT TTTGGGGAG AGCTGTGAAT TCTGAAGATT AGGTGTCTTC TCACCCCAAG 420

CTCCAGAAGT CCAGGCTGAG CCAAAACCAAG CTTCAAGTTG TGCCCTGGACT TGGAGAACCA 480
GGAGGTGAGG GGACTGACTA CTTGAAGATC ACATGGAGGA GGAGTCTGAT CCAGGCCAG 540
GCACCAAGGA AAGGCCATGC AAGGACACAG GGAGAAGGGC AGCTGTCTGT AAGCCAGAAA 600
GAGCCTTAC TAGAAACCA ATCAGCCAGA ACCTTCATCT TGGACTTTCC AGCCTTCAGA 660
5 GATGTGAAA AATAAATTC TGTGTATTA CTTAAAAA

Seq ID No: 163 Protein sequence
Protein Accession #: none found

10

1 11 21 31 41 51
| | | | |
ETLQRQGPQL KREALHTCS SGRTHRLLLL EASRPMVCL FHKPSEGHLE TR

15

Seq ID NO: 164 DNA sequence
Nucleic Acid Accession #: NM_020241.1
Coding sequence: 4-1557 (underlined sequences correspond to start and stop codons)

20

1 11 21 31 41 51
| | | | |
GCCATGCGA CCCCAGGAG GTCCCTCCG CGCCCGGCCC TCCTGCTTCT GCTGCTGCTA 60
CTGGGGGGG CCCACGGCCT CTTTCTCTAG GAGCCGCGCC CGCTTAGCGT GGCCCCCAGG 120
GACTACTGGA ACCACTATCC CGTGTCTGTG GGCAGCGGGC CGGACGCGCT GACCCCCGCA 180
25 GAAGGTGCTG ACACCTCAA CATCCAGCGA GTCTGCGGG TCAACAGGAC GCTGTTCATT 240
GGGAGCAGGG ACAACCTCTA CGCGGTAGAG TTGGAGCCCC CCAAGTCCAC GGAGCTGGGG 300
TACCAGAGGA AGCTGACCTG GAGATCTAAC CCCAGCGACA TAAAGTGTGT TGGATGAAG 360
GGCAACAGAG AGGGCGAGTG TCGAACTTTC GTAAAGGTGC TGCTCCTTGG GGACGAGTCC 420
ACGCTCTTTG TGTGCGGTTC CAACGCTTTC AACCGGTGT GCGCCAACTA CAGCATAGAC 480
30 ACCCTGAGC CGTCCGAGA CAACATCAGC GGTATGGCCC GCTGCCGTA CGACCCCAAG 540
CACGCCAATG TTGCGCTCTT CTCTGACGGG ATGCTCTTCA CAGTACTGT TACCGACTTC 600
CTAGCCATTG ATGCTGTCTT CTACGCGAGC CTCGGGAGCA GGCCCAACCT GCGCACCGTG 660
AAACTGACTT CCAAGTGGTT CAAGAGCCTT TACTTTGTCC ATGGGTGGA GTGGGGCAGC 720
CATGTCTACT TCTTCTCCG GGAGATTGCG ATGGAGTTTA ACTACCTGGA GAAGGTGGTG 780
35 GTGTCCGCG TGGCCCGAGT GTGCAAGAAC GACGTGGGAG GCTCCCCCG CGTGTGGAG 840
AAGCATGGA CGTCTTCTT GAAGCGCGCG CTCACCTGCT CTGTACCCGG AGACTCCCAT 900
TTCTACTTCA ACGTCTGCA GGTCTGTACG GGGTGTGTC GCTCGGGGG CCGGCCCGTG 960
GTCTCTGGCG TTTTCTCCAC GCCCAGCAAC AGCATCCCTG GCTCGGCTGT CTGCGCCTTT 1020
GACCTGACAC AGGTGGCAGC TGTGTTTGA GGCCTCTCC GAGAGCAGAA GTCCCCGAG 1080
40 TCCATCTGGA CCGCGTGCC GGAGGATCAG GTGCTCGAC CCGCGCCCG GTGCTGCGCA 1140
GCCCGCGGGA TGCACTCAA TGCCCTCAGC GCCTTCCGGG ATGACATCCT CAACCTTTGT 1200
AAGACCCACC CTCTGATGGA CGAAGCGGTG CCCTCGCTGG GCCATGCCCC CTGATCCTGT 1260
CGGACCCCTA TGAGGCACCA GCTGACTCGA GTGGCTGTGG ACCTGGGAGC CGGCCCTCTG 1320
45 GCGAACCCAGA CCGTTGTCTT CCGTGGTTCT GAGGCGGGGA CGGTCTCTAA GTTCTCTGTC 1380
CGGCCCAATG CCAGCACCTC AGGGAGCTCT GGGCGTGTGT GTCAAGTGGG CCACGCGTGC 1440
AGGTGTGTGT TCCACGAGCG ACGATCGTGG TGGCCCCAGC GGCCTTGGGG TTGCTGAGC 1500
CGACCTGGGG GCTTCCAGAA GGCCCGGGGG CCTCCGAGGT GCGCGTTAGG AGTTTGAACC 1560
CCGCCCTCTC TGCAGAGGGA AGCGGGGACA ATGCCGGGGT TTCAGGCAGG AGACACGAGG 1620
AGGCCCTGCC CCGAAGTCA ATCGCGAGCA GCTGTCTAAA GGGCTTGGGG GCGCTGGGGG 1680
50 CCGCGAAGGT GGGTGGGGCC CCTCTGTAAA TACGGCCCCA GGGTGTGTAG AGAGTCCCAT 1740
GCCACCCGTC CCCTGTGTAC CTCCCCCTCT TGACCTCCAG CTGACCATGC ATGCCACGTG 1800
G

Seq ID No: 165 Protein sequence
Protein Accession #: NP_064626.1

55

1 11 21 31 41 51
| | | | |
MQTPRASPPR PALLLLLLLL GGAHGLFPPE PPPLSVAPRD YLNHYVPVFG SGPRLTPAR 60
GADDLNIQVR LRVNRTLFIG DRDNLRYVEL EPTSTELRY QRKLTWRSNP SDINVCRMKG 120
KQEGECRNFV KVLRLRDEST LFVCGSNAPN FVCANYSIDT LQPVGDNISG MARCFYDPKH 180
ANVALFSDGM LFTATVDFL AIDAVIYRSL GDRPTLRTVK HDSKWFKEFY FVHAVENGSH 240
VYFFFREIAM EFNYLEKVVV SRVARVCKND VGGSPRVLEK QNTSFLKARL NCSVFGDSHP 300
YFNVLAQVIG VVSLGGRPVV LAVFSTPSNS IPGSAVCAFD LTQVAAPFEG RFREKQSPES 360
65 IWTVPVEDQV PRPRPGCAA PMQYNASSA LPDDIILNVK THPLMDRAVP SLGHAPNLR 420
TLMRHQLTRV AVDVGAGFWG NQTVVFLGSE AGTVLKLFLVR PNASTSGTSG RVCQVGHACR 480
VCVHERRSWV PQRPGRWLSR RWGFPKARGP PRCLGLV

Seq ID NO: 166 DNA sequence
Nucleic Acid Accession #: NM_032108.1
Coding sequence: 39-2705 (underlined sequences correspond to start and stop codons)

70

1 11 21 31 41 51
| | | | |
TCCGAGCGGT CACCTCCTCC TGTCGCTGCG CCCTGCGCAT GCAGACCCCG CGAGCGTCCC 60
CTCCCCGCCC GGCCCTGCTG CTTCCTGCTG TGCTACTGGG GGGCGCCAC GGCTCTTTTC 120

CTGAGGACCC GCGGCGCTT AGCGTGCCG CCGAGGACTA CCGAACCCAT TATCCCGTGT 180
 TTGTGGGCGG GCGGCGCGG AGCGTGACCC CCGCAGAGG TGTGACGAC CTCACATACC 240
 AGCGAGTCTT GCGGCTCAAC AGGACGCTGT TCATTGGGGA CAGGACCAAC CTCTACCGCG 300
 TAGAGCTGGA GCGGCGCAGG TCACGCGAGC TCGGTGACCA GAGGAGCTG ACCTGGAGAT 360
 CTAACCCGAG GACATTAAC GTGTGTGGGA TGAAGGSCAA ACAGGAGGCG GAGTGTGAAA 420
 ACTTCGTAAA GGTGCTGCTC CTTCGGGACG AGTCCAGCT CTTTGTGTGC GGTTCACACG 480
 CCTTCAACCC GGTGTGCGCC AACTACAGCA TAGACACCTC GCAGCCCGTC GGAGACAACA 540
 TCAGCGGTAT GCGGCGCTGC CCGTACGACC CCAAGCAGCG CAATGTTGCC CTCTCTCTG 600
 ACGGATGTCT CTTCACAGCT ACTGTACCG ACTTCCTAGC CATGTATGCT GTCATCTACC 660
 GCAGCTTCGG GACAGGCGCC ACCCTGCGCA CCGTGAACA TGACTCCAG TGGTTCAAAG 720
 AGCCTTACTT TGTCCATGCC GTGAGTGGG GCAGCCATGT CTACTTCTTC TTCCGSGAGA 780
 TTGCGATGGA GTTTAACTAC CTGGAGAAGG TGGTGTGTTC CCGCTGGCC AGAGTGTGCA 840
 AGAACGACTT GGGAGGCTCC CCGCGCTGTC TGGAGAAGCA GTGACCTCC TTCTGAAAG 900
 CCGGCTCAAC CTGCTCTGTA CCGCGAGACT CCGATTCTTA CTTCACCGTG CTGCAAGCTG 960
 GCGAGGCGGT GGTACGCTTC GCGGCGCGGC CCGTGGTCTT GCGGCTTTT TCCACGCCCA 1020
 GCAACAGACT CCGTGGCTCG GCTGTCTGCG CCTTTGACCT GACACAGGTG GCAGCTGTGT 1080
 CCGAGGCGAG CTTCCGAGAG CAGAAGTCCC CCGAGTCCAT CTGACGCGCG GTGCGGAGG 1140
 ATCAGGTGCC TCGACCCCGG CCGCGGTGCT GCGCAGCCCC CCGGATGCGA TACAATGCCT 1200
 CAGCGCGCTT GCGGATGAC ATCCTCAACT TTGTCAAGAC CCACCTCTG ATGACGAGG 1260
 CCGTGCCTCT GCTGGCCCAT GCGGCTTGA TCCTGCGGAC CCGTATGAGG CACCAAGCTA 1320
 CTGAGCTGGG TGTGACGCTG GAGGCGCGCC CCGTGGGCAA CCGACCGCTT GTCTCTCTG 1380
 GTCTCTGAGC GGGAGCGGTC CTCAAGTTCC TCGTCCGCGC CAATGCCAGC ACCTCAGGGA 1440
 GTCCTGGGCT CAGTGTCTTC CTGGAGGAGT TTGAGACCTA CCGGCGGAC AGGTGTGGAC 1500
 GCGGCGCGCG TGGGAGACA GCGCAGCGGC TGTGAGCTT GAGCTGAGC GCAGCTTCGG 1560
 GCGGCTGTCT GGTGCTCTTC CCGCGCTGCG TGTGCGGAGT GCGTGTGGCT CCGTCCGAGC 1620
 AGTACTCGGG GTGTATGAAG AACTGTATCG CAGCTCAGGA CCGCTACTGC GGGTGGGCGC 1680
 CCGAGCGCTC TGTACTCTTC CTGAGCGCGG GCACAGGAGC CCGCTTTGAG CAGGACGTGT 1740
 CCGGCGCGAG CACTCAGGC TTAGGCGACT GCACAGGACT CCGGCGGCGC AGCCTCTCCG 1800
 AGGACCGCGG GGGGCTGCTG TCGGTGAACC TGTGCTAAC GTGCTGCGTG GCGGCTTCG 1860
 TGGTGGGAGC GTGTGTCTTC GGTCTCAGCG TGGGCTGCTT CCGGCGCTC CCGTGGGCGC 1920
 GCGAGCTGCG CCGGCGCAAG GACAGGAGG CCGTCTGCG GCACGCGCGG GCGGAGCGGG 1980
 GCTGTGAGCG CAGCGCGCTG GCGGCGAGCG GCGGCGAGGG TCCGCGGCGC CCGGCGGAG 2040
 GCGGTGCGCG TGGCGCGGG GTTCCCGCGG AGGCGCTGCT GCGGCGCGCT ATGACAGAGC 2100
 GCTGTGCGCA GGCACGCTG CTGAGGCGG GCGGCGCAAG CCGGAGCTG GCGGCTGCTG 2160
 CCAAGCGCGA GACAGCGCG CTGCGCGAGA AGCGCTGCG CACTCCGAC CCGCACCGCC 2220
 ACGCGCTGGG CCGGCGCGCG TGGGACCAAG GCGACCGCTT GCTCCGCGC TCGGCTTCAT 2280
 CCGGCTCTCT GCTGTGCGG CCGGCGCGGG CCGGCGAGCA GCGGCGCGCG CCGGCGGAGC 2340
 GCGGCGCGCA GCGGCGCGCT TATGCTGCGG GCGGCGCGCG CCGCTCCAC GCGGAGCTTC 2400
 CCGGCGCGAG CAGCGCGAG CCGGACCGCG GCGGCGGCTG GTGCGGCGCG CCGGCGCGCT 2460
 TGGACCGCAG CTCAGCGCG CCGGACCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG 2520
 GCGTGGGAGC GCGCTGCGG CCGGACCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG 2580
 CCGTCAACAG CCGGAGCGCG CCGGCGCGGG ACGGCGCGCG CCGGCGCGCG CCGGCGCGCG 2640
 GCGAGGCTTT GCGGCGCGCT CTGCGCTATG GCGGCGCGCG CAGGAGCTG CCGGCGCGCG 2700
 CCGAGCGCGG GCGGCGCGCG ATGCGCTTGG AGTGCGAGCG ACGGAGCAAC GAGGCGAGAG 2760
 AGTGTGCGAG ACGGCGCGCG CCGGCGCGCA CTCGAGTGG GTGCTCAAGT CCGGCGCGCG 2820
 ACCGACCGCG GAGTGGGGG GCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG 2880
 CCGGACCGCG GCGGCGAGGA CCGGCGAGCG GTTGGGCGG GCGGCGCGCG GAGGAGCTTG 2940
 CTATGAGATT GAGGTTGACC TTATGCGCGT AGGTTTGTGT TTTTGTGCA GTTGTGTTT 3000
 CTTTGTGCGT TTTTAAACA ATTGACCAAC TCGGTTCTCG GGTGCGCGCG AGGCGAGGGA 3060
 GCGTGTGAGG CCGGCGCGGA ATGCGGCGCG ACAGCTGCG ACCTAAGCCC TCGGCGCGCG 3120
 CTGGAAGAGT CCGTCCGCAA CCGGCGCGCG TGGGCTGTGT GGTGCTGCGT GCGTGTGCGT 3180
 GCGGCTGTG TGTGCAAGG GCGGCGGAGG TGGGCGTGTG TGTGCTGCG AGCGAAGGCT 3240
 GCTGTGCGCG TGTGTGCAA GTGCGCGAG CCGGCGCGCG GTGCTGCGC GAGCGAGCT 3300
 CCGTGTGCGC CCGGCGCGCT GCGGCGTGGC TGAGCGCGAG CTGCGGCTTC CAGGAGCGCG 3360
 GCGGCTCTCC GAGGTCGCG TTAGGAGTTT GAACCGCGCG CACTCTGCG AGGGAAGCGG 3420
 GACCAATGCC GCGGTTTCAG GCAGGAGACA CAGGAGGCG CTGCGCGAA GTCACATCGG 3480
 CAGCAGCTGT CTAAGGCGCT TGGGCGCGCG GCGGCGCGCG AAG

60 Seq ID No: 167 Protein sequence:
 Protein Accession #: NP_115484.1

1 11 21 31 41 51
 65 MQTPRASPPR PALLLLLLLL GGAHGLFPED PPPLSVAPRD YLNHYPPVFG SGFGRLLTPAE 60
 GADDLNIQVR LRVNRTLPFG DRDNLYRVBL EPPTSTELRY QRKLTVWSNP SDINVCRMKG 120
 KQBGECRNFV KVLRLRDEST LFCVGSNAFN PVCANYSIDT LQPVGDNISS MARCFYDPKH 180
 ANVALFSDGM LFTATVTDPL AIDAVIYRSL GDRPLRLTVK HDSKWFKEPY FVHAVEWGSH 240
 VYFFPREIAM EFNYLEKVVV SRVARVCKND VGGSPRVLEK QWTFPLKARL NCSVPQDSHP 300
 70 YFNVLQAVTG VVSLGGRPVV LAVFTSPNS IPGSAVCAFD LTQVAAVFEG RFREKQSPES 360
 IFTFVPRDQV PRPRPGCAA RGMQYNASSA LPDDILHFKV THPLMDRAVP SLGHAPWILR 420
 TLMRRLQITRV AVDVAGFPWG NQTVVFLGSE AGTVLKLFLVR PNASTSGTSG LSVFLREFET 480
 YRPDRCGRPG GGTGQRLLS LELDAASGGL LAAPFRCCVR VPARCQQYS GCMKICIGSQ 540
 DPCGAPADG SCIFLSPGTR AAFEGQDVGA STSGLGDCGT LLRASLSERD AGLVSVNLLV 600
 75 TSSVAAPVVG AVVGFVVGW FVGLRERREL ARRKDEKAIL AHGAGAVLS VRLGHRRAQ 660
 GPGRGGGGG GAGGVPPBAL LAPLMQNGWA KATLLQGGPH DLSGLLPTP EQTPLPQKRL 720

PTPHPHPHAL GPRAMDHGHP LLPASASSSL LLLAPARAPE QPPAPGEPTP DGRLYAARPG 780
 RASHGDPFLT PHASPDERRV VSAFTGDLDP ASAADGLPRP WSPFPTGSLR RPLGPHAPPA 840
 ATLRRTHTFN SGEARPGDRH RGCHARPSTD LAHLLPYGGA DRTAPPVP

5

Seq ID NO: 168 DNA sequence
 Nucleic Acid Accession #: AW205664
 Coding sequence: 1-135 (underlined sequences correspond to start and stop codons)

10

1	11	21	31	41	51	
CGGCACGAGG	AGAACAGGGG	CCTCTGCCTC	AGTTTGGCCG	GGAGCCAGCC	AGGGCCCATC	60
CTAATTTGGA	GCACAGTCTT	CCCGGTGCCT	AGACATGCCA	AGGCCCTCC	CACGTGGTAC	120
ACCCCTCTCG	<u>TTTGTACCT</u>	GACCACTGT	TTCAAAACGC	AGGTGTTTCT	GGTTTAGAAA	180
CTTGAAGGCG	GGAAATGTGT	TTCTGTCTCT	CTAGGAAGGG	TCTGCTGAGG	ACCAGACCAC	240
GTAAACCTCA	GTGATCTCTG	ACTCAGCTGC	AGCCCTTACC	TGCTCTGTGC	TGATGATCTA	300
TGCATGGCGT	TATGTAGATC	ACGTGCGGCA	GAGACAGCCA	CTGTCTGTGT	TGCGGGTTTT	360
TAAACAGCT	GCCTGGATG	AAACGGAATA	AACCACTGAT	GCTAAAAAAA	AAAAAAAAAA	

20

Seq ID NO: 169 Protein sequence
 Protein Accession #: AW205664

25

1	11	21	31	41	51
RHEENRGLCL	SLPGSQPGPI	LIWSTVFPVP	RHAKAPPTWY	TLVS	

30

Seq ID NO: 170 DNA sequence
 Nucleic Acid Accession #: AB033100
 Coding sequence: 32-2623 (underlined sequences correspond to start and stop codons)

35

1	11	21	31	41	51	
AGGTCTGGGG	TCCTGAGGCT	GCTGGCAGAC	<u>TATGGGTACA</u>	ACGGCCAGCA	CAGCCACGCA	60
GACGCTCTCG	GCAGGCACCC	CATTGAGGG	CCTACAGGGC	AGTGGCACGA	TGGACAGTCG	120
GCACCTCGTC	AGCATCCACT	CCTTCCAGAG	CACTAGCTTG	CATAACAGCA	AGGCCAAGTC	180
CATCATCCCC	AACAAGGTGG	CCCTGTGTGT	GATCACTGAC	AACGCAAGG	AGGAGTTCCA	240
GATCCATGAT	GAGCTGCTCA	AGGCTCATTA	CACGTTGGGC	CGGCTCTCGG	ACAAACACCC	300
TGAGCACTAC	CTGGTGCAAG	GAGCTCAGGC	CTTACCCACG	GGCGCTACT	TCCTGGTGGC	360
GGATCTCACT	GAGAAGATGG	ATGTGCTGGG	CACCGTGGGA	AGCTGTGGGG	CCCCCAACTT	420
CGGGCAGGTG	CAGGTGGGGC	TCACTGTGTT	CGCATGGGGA	CAGCCACGCC	TCTTAGGGTT	480
CAGGCGGGTC	CTCCAGAAAC	TCCAGAAGGA	CGACATAGG	GAGTGTGTCA	TCTTCTGTGT	540
GCGGAGAGAA	WCTGTCTTTT	TCCTGCTGTC	AGATGAGGAC	TTTGTGTCTT	ACACACCTCG	600
AGCAAGCAG	AACTTTCATG	AGAACTTCCA	GGGCTTGGGA	CCCGGGTCC	GGGTGGAGAG	660
CTGGAGCTG	GCCTTCGGA	AAGAGATCCA	CGACTTTGCC	CAGCTGAGCG	AGAACACATA	720
CCATGTGTAC	CATAACACCG	AGGACCTGTG	GGGGGAGCCC	CATGCTGTGG	CCATCCATGG	780
TGAGGAGGAC	TTGCATGTGA	CGGAGGAGGT	GTACAAGCGG	CCCTCTTTCC	TGCAGCCAC	840
CTACAGGTAC	CACCGCTGTC	CCCTGCCCGA	GCAAGGAGT	CCCTGTGAGG	CCCATGTGGA	900
CGCCTTTGTC	AGTGTTCCTC	GGGAGACCCC	CAGCCTGCTG	CAGCTCCGTG	ATGCCCAAGG	960
GCCTCCCCCA	GCCTCGTCT	TCAGCTGCCA	GATGGGCGTG	GGCAGGACCA	ACCTGGGCAT	1020
GGTCTCTGGC	ACCTCATCTC	TGCTTCACCG	CAGTGGGACC	ACCTCCACGC	CAGAGGCTGC	1080
CCCCACGAG	GCCAGGCCCC	TGCTATGGA	GCAGTTCCAG	GTGATCCAGA	GCTTTCTCCG	1140
CATGCTGCCC	CAGGGAAGGA	GGATGGTGGG	AGAGGTGGAC	AGAGCCATCA	CTGCTGTGTC	1200
CGAGTTGCAT	GACCTGAAG	AAGTGTCTTT	GGAAACACAG	AAGAAGTTAG	AGGTATCCCG	1260
ACCGGAGAG	CCAGCCACAG	GAAGCGGCG	CCGACACAGC	GTCTGGCAGA	GGGCGCTGTG	1320
GAGCCTGGAG	CGATACCTCT	ACCTGATCCT	GTTTAACTAC	TACCTTCATG	AGCAGTACCC	1380
GCTGGCCTTT	GCCTCAGTT	TCAGCGCTG	GCTGTGTGCC	CACCTGAGC	TGTACCGCCT	1440
GCCGTGTGAG	CTGAGCTCAG	CAGGCCCTGT	GGCTCCGAGG	GACCTCATCG	CCAGGGGCTC	1500
CCTACGGGAG	GACGATCTGG	TCTCCCCGGA	CGCGCTCAGC	ACTGTCTAGG	AGATGGATGT	1560
GGCCAACTTC	CGGCGGCTGC	CCCGCATGCC	CATCTACGGC	ACGGCCACGC	CCAGCGCCAA	1620
GGCCCTGGGG	AGCATCCTGG	CCTACCTGAC	GGACGCCAAG	AGGAGGCTGC	GGAGGTTGTT	1680
CTGGGTGAGC	CTTCGGGAGG	AGGCCGTGTT	GGAGTGTGAC	GGGCACACCT	ACAGCCTGCG	1740
GTGGCCTGGG	CCCCCTGTGG	CTCTGACCA	GCTGGAGACC	CTGGAGGCC	AGCTGAAGGC	1800
CATCTAAGC	GAGCCTCCCC	CAGGCAAGGA	GGGCCCCCTG	ACCTACAGGT	TCCAGACCTG	1860
CCTTACCATG	CAGGAGTCT	TCAGCCAGCA	CCGAGGGGCC	TGTCTGGCC	TCACCTACCA	1920
CCGATCCCTC	ATGCCGACT	TCTGTGCC	CCGAGAGGAG	GACTTTGACC	AGCTGCTGGA	1980
GGCCCTCGGG	CGCGCCCTCT	CCAAGGACCC	AGGCACTGGC	TTCTGTGTCA	GCTGCCTCAG	2040
CGGCCAGGGC	CGTACCACAA	CTGGATGGT	GGTGGCTGTC	CTGGCCTTCT	GGCACATCCA	2100
AGGCTTCCCC	GAGGTGGGTG	AGGAGGAGCT	CTGTAGTGTG	CCTGATGCCA	GGTCACTAA	2160
GGGTGAATTT	CAGGTAGTAA	TGAAGTGGT	GCACTGTGTA	CCGATGGGC	ACCGTGTGAA	2220
GAGAGAGGTG	GACGAGCGC	TGGACACTGT	CAGCGAGACC	ATGACGCCCA	TGCACATCCA	2280
CCTGCGGGAG	ATCATCATCT	GCACCTACCG	CCAGGCGAAG	GCAGCGAAG	AGGCGCAGGA	2340
AATGCGGAGG	CTGCAGCTGC	GGAGCCTGCA	GTACTTGGAG	CGCTATGTCT	GCCTGATTCT	2400

5 CTTCACGGG TACCTCCACC TGGAGAAGGC CGACTCCTGG CAGAGGCCCT TCAGCACCTG 2460
 GATGCAGGAG GTGGCATCGA AGGCTGGCAT CTACGAGATC CTTAACGAGC TGGGCTTCCC 2520
 CGAGCTGGAG AGCGGGGAGG ACCAGCCCTT CTCCAGGCTG CGCTACCGGT GGCAGGAGCA 2580
 GAGCTGCAGC CTCGAGCCCT CTGCCCCGGA GGACTTGTCT TAGGGGGGCT TACTCCCTGT 2640
 CCCCCACCC ACAGGGCCCC ACGCAGGCCT GGGGTGTCTG AGGTGCTCTT GGTCTGGAGC 2700
 GGCCTGAGG GGTGTGGGCC TTGAAATGAT TCCCCACTT CCTGGAGAGA CTGAGCGGAG 2760
 TTGGAGCCTT TTTTAGAAG AACTTTTAT AGGACAGGGA GACAGCACAG CCATCCCTTG 2820
 CAAACCAACA AGGTGTGTGG CTGACTCCA GGGAGGAGCA CTCACTGGAG TGCTCACAAG 2880
 10 GTGCACACTG CTGTGTGTAC CTTCGAGACA GGCCTGGGTT CAGCTCCAA GGGGCTCACT 2940
 CCCCCAGTTG CCAAACTACT TGGATCTCTC TGCTCTCTTC TCCCTCTCTC CAGATTGGCC 3000
 TGGCAGCCCC TGGCACAGAG CAGACCCGGC CACTGTAGC TCCCACTTC CTACTCTCTG 3060
 CTGCTCTGCC ATTGCCGCTC CCTCTCTGCG TGCCCAAGCA CTGCCCTCGG CGCTCTGGCA 3120
 GCTGTAGGTG GGTGGAGGGG ACAGTGTCTT GATAGATCT ATTATGTGAA AGGCAGCTTC 3180
 ACCACTTTT CTGACTCTC ATGCCCCCAT CTCGACCTG GAGACTTCA GGAATGACAA 3240
 15 CTAACCCAGC CTGCTGGGGC TGGCAGGATG GTGGAGGTTT CTCAGGAGC TGGAGCTTC 3300
 AGGAGCCCC TCTCATGGG AGGAAGAGC TTCCAGGGGG CGAACGAGC ACAGAGGAG 3360
 AGGCTGTCTC CACTTGTCTG GGAACCTGGG CAGGAGGCAC AGAGGAAGCC AAGGCTTGA 3420
 GCTGAGGCTC CCGCGGATC TCTCTCTGTC CCGGAGGCC AGGATGGCT GGTGCCCCA 3480
 CTGCTGTGAG CAGGAGCCCC AAGGAGTGT AGCTGAGGTT GGTGTCTGGG GTGTCTCTCA 3540
 20 TGGACAGTGA GGTGTGCAAG GGTGCACTGA GGTGTGTGG AGGGGATCAC CTGGGTTCCA 3600
 GGCATCTCTT GCTGAGCATC TTTGAGCCTG CCTTCCGCTG GAGCAGAAA AGGCCAGACC 3660
 CTGCTGAGTT AGAGGCTGCT GGGATCACT GTTTCCACAC AGCGGGAAGG CTGCTGGGAA 3720
 CAGTGTGGAG AGAAGTGCCA TGTGTGCGTT GAGCCTTGCA GCTCTCCAG CTGGGAGACT 3780
 25 GTGCTTGTCTG AAACCCAGGA GCTGAACAGT GAGGAGGCTG TCCACTTGC TTGGCTCACT 3840
 GGAACCAAGA AAGCCTGCTT TTGGTTAGGC TGTGTACTT CTGCAGGAAA AAAAAAAG 3900
 GATGTGTCTT TGGTCTGAT ATTTGAAAAG GGGAGGAGC CGAAGTTGTT CCAATTTATC 3960
 CAGTATTGGA AAATATTGGA CCCCCCTGGC TGAATTCTTT TGCAAGACTA CTGTGTCTCT 4020
 GTTCACTACC TTTTCAGGTT TATTGTTTTT ATTTTTCGAT GAATTAAGAC GTTTTAATT 4080
 30 CTTCGAGAC AAGGTCTAGA TGCGGAGTCA GAGATGGGAC TGAATGGGGA GGGATCTCTT 4140
 GTGCTCTCAT GGTGTGCTCT GACTTTCAGC TGTGTGGGA CCACTGGCTG ATCATCAC 4200
 CTCTCTGCTC CAGTTTCCCC ATCTGTAAA TGGAGAGATA ATACTTGCCT ACCTACCTCA 4260
 CRGGGTGTTT GTGAGGATTC ATTTGTGATT TTTTTTTTTT TTTTGTACA GAGCTTTTAA 4320
 GCATTAATAA CAGCTAAATG TG

35 Seq ID No: 171 Protein sequence;
 Protein Accession #: BAA86588.1

1 11 21 31 41 51
 40 MGTASTAAQ TVSAGTFPEG LQSGTMDSR HVSIIHSFQS TSLHNSKAKS IIPNKVAPVV 60
 ITYNCKEEFQ IHDELLKAHY TLGRSDNTP EHYLVQGAQA LPQGRYPLVR DVTEKMDVLG 120
 TVSGCGAPNF RQVQGLTVF GMGQPSLGF RRVLQKLQKD GHRECVIFCV REEVLFLRAD 180
 EDVFSYTPRD KQNLHENLQG LGPGVRVESL ELAIRKEIHD PAQLSENYVH VYHBTEDLWG 240
 45 BPHAVAIHGE DDLHVTREYV KRPLFLQPTY RYHRLPLPEQ GSPLAQLDA FVSVLRETPS 300
 LLQLRDHNGP PFALVFSQGM GVGRNLMGM LGTLILLHRS GTTSQPEARP TQAKPLPMQ 360
 FQVIQSFLRM VFQGRRMVEE VDRATACAS LHDLEVVLE NQKLEGRIP ESQAQSSGR 420
 HSWQRAIWS LERYFYLIIF NYLHEQYPL AFALSPSRWL CAHPELYRLP VTLSSAGPVA 480
 PRDLIARGSL REDDLVSPDA LSTVREMDVA NFRRVPRMPI YGTAAQPSAKA LGSILAYLTD 540
 50 AKRRLKGVVW VSLREZAVLE CDGHTYSLRW PGPVVPDQL ETLAQLKAH LSEPPPKKEG 600
 PLTYRPTCL TMQEVFSQHR RACPLTYHR IPMFDFCAPR EEDFDQLLEA LRAALSXDPG 660
 TGFVFSCLSG QGRITTTAMV AVLAFWHIQ FPEVGEELV SVPDAKFTKG EPQVVMKVQ 720
 LLPDGHVRVK EVDAAALDVS ETMTPMHYHL REIIICTYRQ AGAAKEAQEM RRLQLRSLQY 780
 LERYVCLILF NAYLHLEKAD SWQRPFTWM QEVASKAGIY EILNBLGFFE LESGEDQPPS 840
 55 RLRYRWQBS CSLEPSAPED LL

55 Seq ID NO: 172 DNA sequence
 Nucleic Acid Accession #: AK021806.1
 Coding sequence: 1-645 (underlined sequences correspond to start and stop codons)

60 1 11 21 31 41 51
 65 ACTGTGCTTT TCCTGCGTGC AGATGAGGAC TTTGTGTCTT ACACACCTCG AGACAAGCAG 60
 AACCTTCATG AGAACCTCCA GGGCCTTGGG CCGGGGGTCC GGGTGGAGAG CCTGGAGCTG 120
 GCCATCCGGA AAGAGATCCA CGACTTTGCC CAGCTGAGCG AGAACACATA CCATGTGTAC 180
 CATAACACCG AGGACCTGTG GGGGGAGCCC CATGCTGTGG CCATCCATGG TGAGGACGAC 240
 TTGCTGTGGA CGGAGGAGGT GTACAAGCGG CCCCTCTTCC TGCAAGCCAC CTACAGGTAC 300
 70 CACCGCCTGC CCTGCCCGA GCAAGGGAGT CCCCTGGAGG CCCAGTTGGA CGCCTTTGTC 360
 AGTGTCTTCC GGGAGACCCC CAGCCTGCTG CAGCTCCGTG ATGCCACGG GCTCCCCCA 420
 GCGCTCGTCT TCAGCTGCCA GATGGGCGTG GGCAGGACCA ACCTGGGCAT GGTCTGGGC 480
 ACCCTCATCC TGCTTCACCG CAGTGGGACC ACCTCCAGC CAGAGGCTGC CCCCAGCAG 540
 GCCAAGCCCC TGCTATGGA GCAGTTCCAG GTGATCCAGA GCTTCTCCG CATGGTGGCC 600
 CAGGGAAGGA GGATGGTGGG AGAGGTGGAT AGATCTATTA TGTGAAGGC AGCTTCAACC 660
 75 AGTTTCTTGG ACTCTCATGC CCCCCTCTCC GACCTGGGAG ACTTCAGGAA TGACAACTTA 720
 CCGAGCTGGG TGGGGCTGGC AGGATGGTGG AGGTTTCTCA AGGAGCTGGA GACTTCAGGG 780
 AGCCCTCTCT ATGGGGAGGA AAGAGCTTCC AGGGGGCGAA CCGACGACAG AGGAAGAGGC 840

5 CTGCTCCACT TGCTGGGAA CCTGGGAGG AGGCACAGG GAAGCCAAGG CCTGGAGCTG 900
CAGGTCCCCC GGCACTCTCTC TCTGTCCCGG CAGCCACAGG TGGCTGGTG CCCCACCTG 960
CTGCAAGCAGG AGCCCAAGG AGTGTAGCT GAGGGTGGTT GCTGGGGTGG TCCTCATGGA 1020
CAGTGAAGTG TGCAAGGGTG CACTGAGGGT GGTGGGAGGG GATCACTGG GTTCCAGGCC 1080
ATCCTTGCTG AGCATCTTTG AGCCTGCCTT CCGGTGGAG CAGAAAGGC CAGACCTGTC 1140
TGAGTTAGAG GCTGTGGGA TCCACTGTTT CCACACAGCG GGAAGGCTGC TGGGAACAGG 1200
TGGCAGAGAA GTGCCATGTT TGCCTTGAGC CTTCAGCTTC TTCCAGCTGG GACTGGTGC 1260
TTGCTGAAGC CCAGAGAGCTG AACAGTGAGG AGGCTGTCCA CCTGTCTGG CTCACTGGGA 1320
CCAGSAAAGC CTGTCTTTGG TTAGGCTGCT GTACTCTGTC AGSAAAAA AAAAAGGATG 1380
10 TGTCTTGGT CATGATATTT GAAAGGGGA GGAGGCCGAA GTTGTTCCTA TTTATCCAGT 1440
ATTGAAAAAT ATTTGACCCC CTGGGCTGAA TTCTTTTGCA GAACACTGTT GTGTCTGTTT 1500
ACTACCTTTT CAGGTTTATT GTTTTATT TTGCATGAAT TAAGAGCTTT TAATTCTTTT 1560
GCAGACAAGG TCTAGATGCG GAGTCAGAGA TGGGACTGAA TGGGGAGGGA TCCTTTGTGT 1620
TCTCATGGT GGTCTGACT TTCAGCTGTG TTGGACCAC TGGCTGATCA CATCACTCT 1680
15 TCGCTCAGT TTCCCATCT GTAAAAATGG AGAATAATAC TTGCTACCT ACCTCACGGG 1740
GGTGTGTGA GGAITCATTT GTGATTTTTT TTTTTTTTT TGTACAGAGC TTTTAAGCAT 1800
TAAAAACAGC TAAATGTG

20

Seq ID No: 173 Protein sequence
Protein Accession #: AK021806.1

25 1 11 21 31 41 51
| | | | | |
TVLPLRADED FVSYPTRDKQ NLHENLQGLG PGVRVESLEL AIRKEIHDFQ QLSENTYHVY 60
HNTEDLMGEP HAVAIHGDD LHVTEVYKR PLFLQPTTRY HRLPLPEQGS PLEAQLDAFV 120
SVLRETPSLQ QLRDAHGPPF ALVFSCQMGV GRNLMGVIG TLILLHRSST TSQPEAAPTQ 180
30 AKPLPMBQFQ VIQSPLRMVP QGRRMVEBVD RSIM

35

Seq ID NO: 174 DNA sequence
Nucleic Acid Accession #: NM_016580.2
Coding sequence: 1212-4766 (underlined sequences correspond to start and stop codons)

40 1 11 21 31 41 51
| | | | | |
GGGAAGCGGG AGGAGAGCCA CACGGTCAAG TTGCACAGGT TCTTGACGCT TCTGGAATCA 60
AGACCATGGG CACCTCATA AGTCAGTGTG GGCAGGGACT GCCCAGGGC CAATCCAAGA 120
TCCAGAGGTA GCCATAGGGT GTGACAAGTT GTGCAGATTA CAACACTCAC CCCTTGCAAT 180
AACGTCACTG CCTGTGACTC GGGGCCAGGC CCAGGCCAAA GCCTTCCTTA CATCATTTG 240
TTTAATCCTC ACAGTTTCTT GCTGAAAGGG CTACTATTCT TACTCCATC CCCACTCTAC 300
45 AGATGAGGTA ATGGAGGCCC AGGAAAGTTA AGTGACTTGT CCCAGATGAC ACCGCTGGTA 360
AGTTGCAAGG TCAGAAATTG AACTCAGGCA GTTACCTCT GATGGCTGCT CTGTTAATCA 420
CAGCTGCTTT CCAGTGAGAC AAAAACGGGT GATCAGGGCA GAGTCAAGAC AGAGAGGTAA 480
ACAGATTTGG GAAAGAGACA GGAATGAGAG GGAACAATG GGGGAAAGA TAGGAACAAA 540
GAGAGTTGGG GAAGGGGAGA GAAACAGGAA ACATGACTTG CCGGGAGGG GCATCAGTCC 600
50 ACGTCAAGC AGGTGGAGGC TCAAGTTTTT TGCTCACTTG GTGATGCAGA GGCTCCCTTT 660
CCCTCAGCAG CGGCTTGCT GCGTGGACAG CAGCTTCCA TCTGGCTGT CCGCGGAGCC 720
CGGCTCAT CCTCCTCAGC GGCAGGCCAC TTAGCTTCAC AGGAATGCT CTTTCTCTAA 780
TTGGCATTTA AACTCAGC CCTCCCTTTT CTTGAGGTG GGGTTCCAT AGGAAAGGC 840
TGTCTCTCTG TTTCCCTCAG CTAGCAACTG TTTGGCAGTC AGAGTCCAC ATCTGCTCA 900
55 ACTGGGTGAG GTCCCTCTTA GACCAGCTCT TGTCCATCAT TTGCTGAAGT GGACCAACTA 960
GTTCCCAATG AGATGGCCTT CCCTGGCAAT TCTTGATCGG CATTGGACA TCTCAATCG 1020
GATCCAAATG AGATGGCCTT GCTTGGGGT CTTGCTTUTT TCATAATCAT CTAACATATG 1080
GACAAAGTTG TGCCGGCAGC TCTGGGGGAA GGAGCAGGG GCTGATCAAG CCATCCAGGA 1140
60 AACACTGGAG GACTTGTCCA GCCTTGAAAG AACTCTAGTG GTTCTGAAT CTAGCCCACT 1200
TGGCGGTAAG CATGATGCAA CTTCTGCAAC TTCTGCTGGG GCTTTTGGGG CCAGGTGGCT 1260
ACTTATTCTT TTTAGGGGAT TGTGAGGAGG TGACCACTCT CACGGTGAAA TACCAAGTGT 1320
CAGAGGAAGT GCCATCTGGT ACAGTGATCG GGAAGCTGTC CCAGGAAGTG GCGCGGAGG 1380
AGAGGCGGAG GCAAGCTGGG GCTGCCTTCC AGGTGTTGCA GCTGCCTCAG GCGCTCCCA 1440
TTAGGTGGA CTCTGAGGAA GGCTGTCTCA GCACAGGCAG GCGGCTGGAT CGAGAGCAGC 1500
65 TGTGCGSACA GTGGGATCCC TGCTGTGTTT CTTTGTATGT GCTTGCCACA GGGGATTGG 1560
CTCTGATCCA TGTGAGATC CAAGTGCTGG ACATCAATGA CCACAGCCA CGGTTTCCA 1620
AAGCGAGACA GAGCTGGA ATCTCTGAGA GCGCTCTCT GCGAACCGG ATCCCTCTG 1680
ACAGAGCTCT TGACCCAGAC ACAGGCCCTA ACACCTGCA CACCTACACT CTGCTCCCA 1740
GTGAGCACTT TGCTTTGGT GTCATTTGG GCGCTGATGA GACCAACAT GCAGAACTCA 1800
70 TAGTGTGAA GGAGCTGGAC AGGGAATCC ATTCATTTT TGATCTGGTG TTAACCTGCT 1860
ATGACAAATG GAACCCCCC AAGTCAGGTA CCAGCTTGGT CAAGGTCAAC GTCTTGGACT 1920
CCAATGACAA TAGCCCTGCG TTTGCTGAGA GTTCACTGGC ACTGGAATC CAAGAAGATG 1980
CTGCACTGGT TAGCTTCTC ATAAACTGA CCGCCACAGA CCCTGACCA GGCCCAATG 2040
GGGAGGTGGA GTTCTTCTC AGTAAGCACA TGCCCTCAGA GGTGCTGAG ACCTTCACTA 2100
75 TTGATGCCAA GACAGGCCAG GTCAATCTGC GTGCACTCT AGACTATGAA AAGAACCTG 2160
CCTACAGAGT GGATGTTACG GCAAGGGACC TGGTCCCAA TCCTATCCCA GCCCATTGCA 2220

5
 10
 15
 20
 25
 30
 35
 40
 45
 50

AAGTCTCTCAT CAAGGTTCTG GATGTCATG ACAACATCCC AAGCATCCAC GTCACATGGG 2280
 CCTCCAGCC ATCACTGTTG TCAGAAGCTC TTCCCAAGGA CAGTTTATT GCTCTGTCA 2340
 TGGCAGATGA CTGGATTC GAACACATG GTTGTGTTCA CTGCTGGCTG AGCCAAGAGC 2400
 TGGGCCACTT CAGGCTGAAA AGAACTAATG GCAACACATA CATGTTGCTA ACCAATGCCA 2460
 CACTGGACAG AGAGCAGTGG CCCAATATA CCTCACTCT GTTAGCCCAA GACCAAGGAC 2520
 TCCAGCCCTT ATCAGCCAA GAAACAGCTCA GCATTCAAGT CAGTGACATC AACGACAATG 2580
 CACCTGTGTT TGAGAAAAG AGGTATGAAG TCTCCACGGG GGAAAACAA TTAACCTCTC 2640
 TTCACCTCAT TACCATCAAG GCTCATGATG CAGACTTGGG CATTAAATGA AAGTCTCAT 2700
 ACCGCAATCA GGAATCCCA GTTGCTCACT TAGTAGCTAT TGACTCCAA CAGGAGAGG 2760
 TCATGCTCA GAGTCACTG AACTATGAAG AGATGGCGG CTTTGAGTTC CAGGTGATG 2820
 CAGAGGACAG CGGCAACCC ATGCTTGCAT CCAGTGTCTC TGTGTGGGT AGCTCTTGG 2880
 ATGCAATGA TAATGCCCA GAGTGTGTC AGCTGTGCT CAGCGATGGA AAAGCCAGCC 2940
 TCTCCGTGCT TGTGAATGCC TCCACAGGCC ACCTGCTGTT GCCCATGAG ACTCCCAATG 3000
 GCTTGGGCC AGGGGGCACT GACACACCTC CACTGSCAC TCACAGCTCC CGGCCATTCC 3060
 15 TTTTGAACAC CATTGTGGCA AGAGATGCAG ACTCGGGGGG AAATGGAGAG CCCCTCTACA 3120
 GCATCCGAG TGAATATGAA GCCCACTCT TCATCTCAA CCTCATACG GGGCAGCTGT 3180
 TCGTCAATGT CACCAATGCC AGCAGCTCA TTGGGAGTGA GTGGGAGCTG GAGATAGTAG 3240
 TAGAGGACCA GGAAGCCCC CCTTACAGA CCGAGCCCT GTTGAGGCTC ATGTTTGTCA 3300
 CCAGTGTGGA CCACTGAGG GACTCAGCCC GCAAGCCTGG GGCTTGAGC ATGTGATGC 3360
 20 TAGCGTGTAT CTGCTGGCT GTACTGTGG GCATCTTGG GTTGATCCTG GCTTTGTTC 3420
 TGTCCATCTG CCGACAGAA AAGAAGGACA ACAGGGCTTA CAACTGTGG GAGGCGAGT 3480
 CCACCTACCG CAGCAGCCC AAGAGGCCCC AGAAACACAT TCAGAAAGGA GACATCCACC 3540
 TCGTGCCTGT GCTCAGGGT CAGCAGGTG AGCTTGTGA AGTCGGCG AGCCACAAAG 3600
 25 ATGTGACAAA GGAGGAGATG ATGGAAGCAG GCTGGGACCC CTGCTGAG AGCCCTTCC 3660
 ACCTCACCCC GACCTGTGAC AGGACGTGTC GTAATCAAGG CAACAGGGA GCACCGCGG 3720
 AGAGCCGAGA GTGTCTGCAA GACACGTGTA ACCTCCTTTT CAACATCCC AGGCAGAGGA 3780
 ATGCTTCCCG GGAGAACCTG AACCTTCCCG AGCCCAAGCC TGCCACAGGC CAGCCAGCTT 3840
 CCAGGCTCT GAAGTTGCA GGCAGCCCCA CAGGAGGCT GGCTGGAGC CAGGCGAGT 3900
 30 AGGAAGCCCC ACAGAGGCCA CCAGCCTCCT CTGCAACCTC GAGACGGCAG CGACATCTCA 3960
 ATGGCAAGGT GTCCCTGAG AAGAATCAG GGCCCGTCA GATCCTGCG AGCTGTGTC 4020
 GCTGTCTGT GGCTGCTTC GCGAGCGGA ACCCGTGGG GAGGCTCACT GTGATTTCT 4080
 CTCTGTGTA GCAATCTCC CAGCTGCTGT CCTTGTGCA TCAGGGCCAA TTCCAGCCCA 4140
 AACCAACCA CAGAGAAAT AAGTACTTGG CCAAGCCAGG AGGCAGCAGG AGTGCATCC 4200
 35 CAGACACAGA TGCCCAAGT GCAAGGCTG GAGGCCAGC AGACCCAGAA CAGGAGGAG 4260
 GGCCTTTGGA TCCTGAAGAG GACCTCTCTG TGAAGCAACT GCTAGAAGAA GAGCTGTCAA 4320
 GTCTGTGGA CCCAGCACA GGTCTGCCCC TGAGACGGCT GAGCGCCCT GACCCGCTC 4380
 GGATGGCAG ACTCTTTT CCCTCACCA CCAACTACCG TGACAATGTG ATCTCCCGG 4440
 40 ATGCTGCAGC CAGGAGGAG CCAAGGACCT TCCAGACCTT CGGCAAGGCA GAGGCACAG 4500
 AGCTGAGCCC AACAGGACG AGGCTGGCCA GCACCTTTGT CTGGAGATG AGCTCACTGC 4560
 TGGAGATGCT GCTGGAACAG CGCTCCAGCA TGCCCGTGGG GGCCGCTCC GAGGCGCTG 4620
 GGCGGCTCTC GGTCTGCGGG AGGACCTCA GTTTAGACTT GGCCACCACT GCAGCCTCAG 4680
 GCATGAAGT GCAAGGGGAC CAGGTGGAA AGACGGGAC TGAGGGCAAG AGCAGAGGCA 4740
 45 GAGCTGAGC CAGCAGGTG CTGTGAACAT ACCTCAGACG CCTCTGATC CAAGAACCAG 4800
 GGGCCTGAGG ATCTGTGAG AAGAGCTGGT TTCTAAATC TTGTAACTCA CTAGCTAGC 4860
 GAGGCTGAG AACTTAGGG TGAATGATG TACCCCAACA GAGGAGGCAA GAGCCCAAG 4920
 ACTAACAGT GACTACCAA AGCAGCCCTT TGTAAAGCAG TCTGAGTCTT TTGAGGACA 4980
 GGGACGTTT GTGCTGAGA TAAGTGTTC CTGGCAAAAC ATATGTGAG CACAAAGGT 5040
 CAGCTCTCTG GCAGAACAGA TGCCACGAG TATCACAGGC AGGAAAGGT GGCCTCTCTG 5100
 50 GGTAGCAGA GTACGGGGG GTTACCTTG GGTGCCAGG AATGCTCTC TGACCTATCA 5160
 ATAAAGGAAA AGCAGTATT CAAAAA AAAA AAAA

Seq ID No: 175 Protein sequence:
 Protein Accession #: NP_057664.1

55
 60
 65
 70
 75

1 11 21 31 41 51
 NMQLQLQLL LIPQGVLP LQDCQEVTL TVKYQVSEV PSQTVIGKL QELGRZERRR 60
 QAGAAQVQL LPQALPIQVD SEBGLLSTGR RLDRBQLCRQ NDPCLVSFV LATGDLALIH 120
 60 VEIQVLDIND HQPRFPKGEQ ELBISSEASL RTRIFLDRL DPTGPNLTH TTLSFSEHF 180
 ALDVIVGDE TKHAELIVK ELDRBIHSFF DLVLTAIDNG NPPKSGTSLV KVMVLDSDND 240
 SPAPABSSLA LEIQEDAAPG TLLIKLTATD PDQFPNGEVE FFLSKHMPPE VLDTFSIDAK 300
 TGGVILRRPL DYKKNPAYEV DVQARDLQPN PIPAHCKVLI KVLVDVNDIP SIHTWASQP 360
 65 SLVSEALPKD SFIALVMADD LQSGHNLVH CWLSQBLGHP RLKRTNGNTY MLLTNATLDR 420
 EQWPKYTLFL LAQDGLQPL SAKKQLSIQI SDINDNAPVF EKSRYEVSTR ENNLPSLHLI 480
 TIKAHADLNG INKVSRIQ DSVPAHLVAI DSMTEVTAQ RSLNRYEMAG FEFQVIAEDS 540
 GQPMADSSVS VMVSLDAND NAPEVVPVVL SDGKASLSVL VNASTGHLLV PIETPNGLGP 600
 AGTDTPLPAT HSSRPFLITT IVARDADSGA NGEPFLYSIRS GNEAHLFILN PHTGQLFVNV 660
 70 TNASSLIGSE WELBIVVEDQ GSPPLQTRAL LRVMFVTSVD HLRDSARKPG ALSHSLMTVI 720
 CLAVLLGIFP LILALFMSIC RTEKKDNRAY NCREABSTYR QQPKRPQKHI QKADIHLVFPV 780
 LRQGAGBPCE VQSHKQVDK EAMMEAGWDP CLQAPFHLTP TLYRTLNGG NQAPABSSRE 840
 VLQDTVNLFP MHPQRNASR ENLNLPRPQP ATGQPRSRPL KVAGSPTRGL AGDQSEBAP 900
 QRPPASSATL RRQRHLANGV SPEKESGPRQ ILRLSVRLSV AAFARNPVE ELTVDSPPVQ 960
 75 QISQLSLHLH QQQPQPKPNH RGNKYLAQPG GSRSAIPDTP GPSARAGGQT DPEQSEBQPLD 1020
 PEEDLSVQKL LBEELSLILD PSTGLALDRL SAPDPANMAR LSLPLTNYR DNVSFPAAD 1080
 TERPRTFQTF GKAEAPELSP TGTRLASTFV SEMSLLLEML LEQRSSMPVE AASEALRRLS 1140

VCGRTLSDL ATSAASGMKV QGDPGGKTGT EGKSRGSSSS SRCL

Seq ID NO: 176 DNA sequence

Nucleic Acid Accession #: AL109712.1

5 Coding sequence: 2-128 (underlined sequences correspond to start and stop codons)

1	11	21	31	41	51	
10	GAGTCTCTTT	GGGCCAGCG	GGCTGCTGCA	GACAGACAGG	AAGCAGCCT	GACGCTCTTC 60
	TACCCTCGGG	CAGCACAGCG	GGGCTGGGAC	TCACTCTAGC	TGCCCCAGCA	ACTTGCTTTC 120
	CTGTGTGAC	TCTGGCAGGC	TGCCCTCTCT	GTGCAAGCT	GCCACTGGGG	CCTGCTCAGG 180
	GTGGCTGGA	ACTTGGAGGT	GGGCACTCAG	GGCTTAGGAT	GGGCTGTGT	CACCAAGGCA 240
	TGTGCCCTTG	GGCCAGTTAC	TTCTCTCAG	AGCCTTGGGC	TCCTCTCTGT	AGGATGGGGC 300
15	TTGTTGGTGT	GAAATGAGGT	GAGCATGTTG	AGTTGGGGAG	CAGCAGGACA	CGCACCTGCA 360
	GGCAGCGGCC	CTGGCCACGC	TCCCTCCCTA	CCTTCCGAGT	CCTGGGACAG	ACACAGTAGA 420
	GCACAGCGGG	CCAGCCTGCT	CTCTTCTCTG	TCTACTTTTT	GCAGAAGAGT	CAACAGATAC 480
	AACAGGCCCA	GGGAGGTGCC	CCTGGGGGCC	CCAGTCCCCA	TCACTCCAAG	GGGCAGTCTC 540
	GCAGGTGACA	AGGTGGGCC	AATCCCTGTG	GAACAGGTCT	CTGAGGACCA	CAGAGTGGGG 600
20	CCCCAGGAA	AGCTGGGAGC	CGAGCTAGAG	CGAGGCAGCA	AGTAAGGCCA	AAGCTGTGCC 660
	CCTGCCCGGA	AGAAGCTTCT	GCCCCCAGAA	CCGACCCCTC	CGCAGATAGC	CCTCCCTGGG 720
	CAGCAGCGCC	CCAGCTTCCA	AGGCCGCTGC	CTCACCAGAC	GCCATGCTCT	CACGACTTGT 780
	TTTGCTGTCT	TGTACCTGTC	AGATCTGCCC	CAGAGGAGCA	GGTGAAGAGC	CGCGCTTGCC 840
	AGGCTCTGT	GGCGTGGAG	TTTTGGGCAG	AGGAGTGGGG	GGAGAGTTT	CTCATTTTA 900
25	AGATTCTCCA	AATCAAGAT	GAACTCATGC	TGTGCTTTGG	AATGATGAT	GCTCATTTTA 960
	GTAAATCAT	AATAATGTT	ACACAACTG	TTAAAAAAA	AAAAAAA	AAAAA

Seq ID No: 177 Protein sequence:

Protein Accession #: AL109712.1

30	1	11	21	31	41	51
	VSLQPGGCCR	QTGSTPDAPL	PSGSTAGLGL	TLACPATCFP	V	

35 Seq ID NO: 178 DNA sequence

Nucleic Acid Accession #: none found

Coding sequence: 3-107 (underlined sequences correspond to start and stop codons)

40	1	11	21	31	41	51	
	AATGGAGCAC	TCGAAAGAAC	GATTTGACCA	ATAGCATTTC	TTCTCTGGGG	GTGTATTTC	60
	AAAGCATGCA	ACTCTCCAGG	GAACCAAGAC	TAAATTGCTT	AAAATGAAGT	CATTCTCTAG	120
	ATTAACTTCC	TCAGATAAAG	TGTCAAGCGT	CTGCAGAAAC	GAAGAAGACA	AAACTGAGAT	180
45	TATCACTCAT	AATTCTCTTA	CTTACTATGT	CAGTGAAGCA	ATGAGTTTGC	ATTTTTCACA	240
	TCCTAGAACCA	TTCTTCAITTA	GCCTGGGTC	ATGACCTCTT	CCAGTTAATT	CTCTTTCACA	300
	CCTTTAGGAA	AGATTTAAGA	TGAACCTTCA	ATAGGATATT	AACATAACTC	ATAGCCCAATA	360
	CCACAGCTGC	CTTTCAAATT	AATGAGTTTA	ATTGTTCTCC	AGCAAACATG	AGTTTGTCTT	420
	TGGCATTTTA	AATGCTTCCC	ATTGATCTGA	CATTTTGCTG	TTTCAAGTTT	TAAAGGGCTC	480
50	AAATCAAGCA	CTATTGATAA	CTGAGCAAGG	AGCGAAGATC	CAGAAATACG	AAAACATTGT	540
	CTTTTTTTTT	CCATGAAAAA	CAATCATAGC	CTTTTGAAAT	CAATCGAAGT	TTCTACATTA	600
	GCCATCTAAG	ACTTATTTAA	TTATTCTGT	TCTCAGTCAA	GCTAATTCAA	GTGAATGAAC	660
	AGTATTGACT	TTTAAATCT	TTTTTAAATT	TTTTTAAATC	TTTAGTTTAT	TAAATTGTGA	720
	GAAAAGCTCT	GGGGCCATGA	CCACTTACGT	AAATGTTTCA	GTTTAAAAAC	AAAAGATTCA	780
55	GGCCTCTAAT	TTGAGCCAAA	TCCAGTGTAT	CTTGTTTGAA	ATTTTGTATG	AATTGTAAAA	840
	GATGAAAGTG	GAACTTTTAA	CATTCAATTT	CCCCAAATTT	TTCAGTGGGA	AGGGATGCTA	900
	ATTGCCCTACT	TAAGATATAA	GTTCAGGAAT	AACATTTTCA	TAGAAAAATC	AGAAAACTGC	960
	TTGACACAGC	AGTGACATAG	TTAGATGTTG	CTCAGATGCC	TTCCAAACCT	GAGGGTCCCC	1020
60	AAAGATTCTT	TTACCAATTG	TTTTTAACCTA	TGAATCTTAA	TCTTGTTCAT	TCCCTGCCCA	1080
	AAACAAATTT	AAAAAG					

Seq ID No: 179 Protein sequence:

Protein Accession #: none found

65	1	11	21	31	41	51
	WSTPKNDLTN	SISSLGVVFG	SMQLSRPEL	NCLK		

70 Seq ID NO: 180 DNA sequence

Nucleic Acid Accession #: none found

Coding sequence: 2-176 (underlined sequences correspond to start and stop codons)

75	1	11	21	31	41	51

5 CCGGGTGGGG CCTGGGATG CAGGCGCCGG TCGCCGCGCC CCTGGGCTG CTGACCCCG 60
 CAGAAGGGCT TTCGAGGAGG AAGAAGACGT CGCTCTGGTT TGTGGGGTCT CTGCTGCTGG 120
 TGTCCGTCTT CATAGTCACC GTCGGGCTGG CTGCATCAGC AGGACGGAGA ATGTGACCGT 180
 TGGGGGCTAC TACCCAGGGA TCATTCTCGG CTTTGGATCT TTCTTAGGAA TTATTGGCAT 240
 CAACTTGGTG GAGAAATAGAA GGCAAATGCT GGTGGCAGCG ATCGTGTTTA TCAGTTTGG 300
 CGTGGTGGCC GCCTTCTGCT GCGCCATCGT GGACGGCGTA TTGCGAGCAC AGCACATTGA 360
 ACCGAGGCCC CTCACACGGG GAAGATGCCA GTTTTACTCC AGTGGGGTGG GGTACTTGT 420
 CGATGTCTAC CAGACAGAGG TGAGCAGGAG CACTGAGATT CATGTGGGTT TTGCTCAGCT 480
 AACCCCGCCG ACCCCACGGG GTTTTCCCTG CACATAGGCG TGGTCTGAAT ATTGGATTTC 540
 10 TAATAGTTCC TGGGGGTAC CCTTGCAGCT GGTGAACCGT TGATGCCCCC TGTGTAAAGG 600
 ACCTTGACAT TTGATGTGCA TGTATTTCAC TCTGAGTCA GAGTCTTGA CTGTCTTCAT 660
 TAAATCACAA CAGTCTCAGA AAACAACCGC ACCACCCCGC AATCCACCA AAGGGGCGCG 720
 CCGTCCCTAA GAGTTATCCC

15 Seq ID No: 181 Protein sequence
 Protein Accession #: none found

20 1 11 21 31 41 51
 RVGPRDAGAG ARAPGPAGPR RRAFEEDVDV ALVCGVSAAG VRPHSHRRAG CISRTENV

25 Seq ID NO: 182 DNA sequence
 Nucleic Acid Accession #: AK001579.1
 Coding sequence: 1150-2637 (underlined sequences correspond to start and stop codons)

30 1 11 21 31 41 51
 TTTTCTCTGC TTTTCGCTAC CCGGTCACCT CTCAATTTCT TCCCTATTTC CTGTCTCTTT 60
 CCCCACATCCC CCTTTCTCCT GTCTCTCCCC TGCCCTCTACA GTGGTTCTCC CGCTGAGCT 120
 GCCACACAGT GCTGGGCCCC GGGCTGCTGC GGTCTGGGCG CCTATGGCTG CGGTCCCCT 180
 CCATATCAGC CCGGGCCCCCT GGTCTCTGCG TGTGAGGGTT TGGCTCCTTT CGTGGTGACC 240
 ACCTCTCTCT GTGCTCAGCG CCGGGCCCCAG GCCCCCAGC CCCTGAGGAC ATGTGTGATC 300
 TGCGGCGCTC ACAGAGATGC AGTGTGGTTT CTGCAGCTGA CACCCAGAT AAGAAGAGAG 360
 35 ATTTGCTCCT GGTGGAGACA GGAAGGACCC TGTATCTGCA AGGAGAGGGC CGGTGGAGCT 420
 TCACGGCATG GAACGCAGCC ATTGGGGGCG CGGCTGGTGG GGGCGGCACA GGGCTGCAGG 480
 AGCAGCAGAT GAGCGGGGGT GACATCCCCA TCATCGTGA TGCTGCTATC AGTTTGTGTA 540
 CCCAGATGCG GCTCCGGCTG GAAGGTGTAT ACCGGAAGG GGGCGCTGCT GCCCGCAGCC 600
 TGAGATCCTT GGTCTGAGTT CGTGGGATG CCGGTCTGGT GAGCTCCGGA CCAGGGGAGC 660
 40 ACTTGTGGA GGAATGCTAC GACACACTCA AAGCTCTCT TCGTGGCTC GATGACCTGT 720
 TGACCTCTGC ACGGTGCTG CCTGCTGGA GGGAGGCTGC TGTATTCTCT AAGATCCCTG 780
 AGAGCCAAAG CCCAACGAG ATCTCTGCTT TCCCCACCA GAATCCATGG TTTGGCAGCC 840
 CTCGCCCCCA TCATCTCCCA CCTGGGGGA TCATCCAGAG ACTTGGCTCA GGGGAGAGTG 900
 GGAAGGGGGC AGAGACACAT CCATCTGCA TTTGTGCTA AAATCCCTC CCTCTGTACC 960
 45 AGCTGCCACT CTTTCTTCCC GGTCTCTCCC CAACCTCTCT CCAATCCATC CCCAGAGCTG 1020
 CCCCAGAAAG ATCAGCGCCT GGAGAAATAT AAAGATGTA TTGGTGCCT GCCCGGGTCT 1080
 AACCGCCGCA CACTGGCCAC CTTCTATGGT CATCTCTATC GGGTGCAGAA ATGTGCGGCT 1140
 CTAACACAGA TGTGACGCG GAACCTGGCT CTGCTGTTTG CACCCAGCGT GTTCCAGAGC 1200
 GATGGGCGAG GGGAGCAGCA GGTGCGAGTG CTGCAGAGC TCATGTATGG CTACATCTCT 1260
 50 GTCTTTGATA TGGATTCTGA CCAGGTAGCT CAGATTGACT TGGAGGTGAG TCTTATCACC 1320
 ACCGTGAAGG ACGTGCAGCT GTCTCAGGCT GGAGACCTCA TCATGGAAGT TTATATAGAG 1380
 CAGCAGCTCC CAGACAACTG TGTACCCCTG AAGGTGTCCC CAACCTGAC TGTGAGGAG 1440
 CTGACTAACC AGTACTGGA GATGCGGGGG ACAGCAGCTG GATGGAAGTT GTGGGTGACT 1500
 TTTGAGATTG GCGAGCATGG GGAGCTGGAG CGGCCACTGC ATCCCAAGGA AAAGGTCTTA 1560
 55 GAGCAGGCTT TACAATGGTG CCAGCTCCCA GAGCCCTGCT CAGCTTCCCT GCTCTGTAAG 1620
 AAAGTCCCCC TGGCCCAAGC TGGCTGCTTC TTACAGGTA TCCGAGTGA GAGCCACCG 1680
 GTGGGGCTGT TGGGTGTGCG TGAGGAGCCA CCTGCTGCTG TGGGAAGCCG CTTCAGGAG 1740
 AGTGTCTTTC TGTGCGTGG CCGCTGCTGG CTGCTGCTCA AGGAGAAGAA AAGCTCTAAA 1800
 60 CCAGAACGGG AGTGGCCTTT GGAAGGTGCC AAGTCTTACC TGGGAATCCG CAAGAAGTTA 1860
 AAGCCCCCAA CACCGTGGGG CTTCACATTG ATACTAGAGA AGATGCACCT CTACTTGTCC 1920
 TGCATGTAGC AGGATGAAAT GTGGGATTGG ACCACGAGCA TCCTTAAAGC CCAGCAGGAT 1980
 GACCAAGAGC CAGTGTGCTT ACAGAGCCAT TCTCTCTCTG ACCTTGGCCG TCAGAGGTTT 2040
 GGCATATGCT CTTTGTGCTC TATCGGTGGG GATGACAGTG GAGCCACCTC CCTCTCTGCC 2100
 65 AATCAGACCC TCGGCGACT ACACAAACCG AGGACCTGT CCAATGTTCT TCCAATGAAG 2160
 TCATCCAGG GGTCTGTGGA GGAGCAAGAG GAGCTGGAG AGCCTGTGTA CGAGGAGCCA 2220
 GTGTATAGG AAGTAGGGCG CTTCCTGAG TTGATCCAG ACACCTCTAC CTCTCTCTCC 2280
 ACCACACGGG AGTGGACAGT GAAGCCAGAG AACCCCTCA CCAGCCAGAA GTCAATTGGAT 2340
 CAACCTCTTC TCTCAAGTC AAGCACCCTT GGCCAGGAGG AGAGGCCACC TGAGCCCCCT 2400
 CCAGGGCCCC CTTCAGAGAG CAGTCCCCAG CACGGGGGT CCTAGAGGA ACAGCTGCTC 2460
 70 CAGAGCTCA GCAGCTCAT CTTGAGGAAA GGAGAGCCA CTGCAGGCTT GGAAGTCTCT 2520
 TCCAGGCCAT CCAGCCCCCA ATCCCCCAGC CCCACTGGCC TTCCAAACCA GACACTGGC 2580
 TTCCCCACCC AACCCCCATG CACTTCCAGT CCACCTCCA CCAGGCCCTT CACATGAGCC 2640
 TAGGACAGC AGTCTGAGAG GGTAGGTACC AAGAAGCCCA GAAACTCTTA TCGTGGCACT 2700
 GTTGCAGCTT CTTCTGCCCT GGTGGAAG ACTCCAGAA CAGTGTGGT GCTGTGGAAG 2760
 75 GAGCACTGGA CTAAGGCTT CAGTGGCTGC GTGTCCAGG ACAGGTCAAT GCCCTCTCT 2820
 GGGCCAGACC CATTATCTA TACCATGAG TAACGAAAT AAGGAGAGCA GTGAATGTCA 2880

AACTGTGTTT CTTAGAGCCA TAAGCCCCAC ATATTATCCC TGAACAAGGG CAGCTCCTGC 2940
 TTTATATATT TGATACGTAG GGGTTCCATG AGAGATTITG GGTITTTAAG GAATGTTTTT 3000
 ACTGCATTAA AGAAAAAAA TGCTTTGGAA ACCAGAGGCC TGGGTGATGT TAAAGTCTAT 3060
 CCTGTCCAC TCTCTACATT CTGGGACTAC CGTGAAGCCT GGAOTAGGGA GAGCGAGTTT 3120
 GGGAGCTGGG ACTCGGGGAG TCAAAAATAG ATGAGTAATT GTCAATAAAC CTGGGAACC

Seq ID No: 183 Protein sequence:
 Protein Accession #: AK001579.1

10 1 11 21 31 41 51
 MSLTHSNASF VSSMTLPLHG CCLAGGRLLV FLRLSLAKAQ PGSLSPTRI HGLAALRPIT 60
 SHPGSSRDAL AGDEVGRQR HIHPAFVPMK PSLCTSCHSP FPGPPQSSI PSPPLQKMQ 120
 15 RLEKRVLDVIG CLPRVNRRL ATLIGHLYRV QKCAALNQM TRLALLFAP SVFQTDGRGE 180
 HSEVRLQELI DGYISVFDID SDQVAQIDLE VSLITWEDV QLSQAGDLIM EYVIEQQLPD 240
 NCVTLLKVSPT LTABELNQV LEMRGTAAGM DLWVTFEIRE HSELERPLHP KEKVLEQALQ 300
 WCQLPBFCSA SLLLLKVPLA QAGCLFTGIR RESPRVGLLR CREBPPRLIG SRPQSRFFLL 360
 20 RGRCLLLLE KSSSKPEREW PLBGAKVYLG IRKRLKPTP WGTLLILEKM HLYLSCTDRE 420
 EMMWTTTSL KQHDHQQPV VLRHSSSDI ARQKFGTMDL LPIRGDSDGA TLISANQTLR 480
 RLMNRRLTLM FPFMKSSQGS VERQEELREP VYBPFVYEV GAFPELIQDT STSPSTTREW 540
 TVKPNPLTS QKSLDQFFLS KSSTLQREER PPEPPPGPPS KSSPQARGLS HBQLLQELSS 600
 LILRGGETTA GLGSPSQSS PQSPSPTGLP TQTPGFPTQP PCTSPSPSQ PLT

25 Seq ID NO: 184 DNA sequence
 Nucleic Acid Accession #: none found
 Coding sequence: 1-81 (underlined sequences correspond to start and stop codons)

30 1 11 21 31 41 51
 GTAGAGTTAG TGTCATGTG CTTAGATAT ACCAAATTCA TAAACATTTT CTCTAAAAA 60
 GTATTAGCT TAAGAAAGTTA ATTCACTTTA AGGAATATAA ACCAAATTAT TTTATATTG 120
 35 AATCTCAACA TAAGAAGTCA AATGTAAATG CTGCCAGATA ACAATATCAA AGGTATTTTT 180
 TTTCTCTAT AATTTCATCA GTATGTCCTC TCCCTTTCT CTTATTGTG AATTTTAGC 240
 AACCTTAAT CTGCTAATTA TAAGCTAGGC AAGTAATCTT GGACAAGTTA TTGACCTCT 300
 CACTGCACCA GCTTTGTAT CTGTAAATG ATGATAATAC CAACACCTTC TTCTGGGGT 360
 ACTGAAGATG AGAGAACATG ATATGTGTAA AGTGCCCTCC ACAATACCCA GAACATAGCA 420
 AACATGTAAT GAATGTAGTA ATAGTAATTA TTTTATTTTC TTTTGATTCA GTTGGGACTA 480
 40 TGTTCAGCTG TAACAGAATA CCCAAATATA CAGTTTAAAA CAATTAAG TTTTGTGTG 540
 AAGTTTGTG ACGAATTGAG ACAATCCAGG GCTTTTATAG ATGCACCAGG ATCAGCAGGT 600
 ACAGAGGCACT CTTTCTGAT TTCTGCCAGT CTCAATGCAT GGGTTGCAAT CCAGAGTCCA 660
 GGATGTCAGT TCCAGCCCTG GTTACGCCCA TATTAGCACA CAGAAAGAAA GAGAAAGGGA 720
 TGTGCTCTCT CACTTTAATC ATAGCTCCCA CTAGATGCAC CCATCTACTC TGCTGATACT 780
 45 CCATAGCTTC ATGCTTGCCT ACATGGTCAC ACTTAGTTTC CAGAGAGACA TGTCTGACA 840
 GTCATGTGCT CAATTAATAT CCAAGTGTCC AATTACTGAG AAAAAAGAA ACTAGCACCT 900
 TTGCTTGTGT GCATTCTCTT TAGCATAAGC CACATCTTTT TTATGAAGTT GTCTCTAGTT 960
 ACTTGATGTC CTCAGTTGTC CTTTCATTTA GAATGCTGCC TTGCACATCC TGAATCTGAC 1020
 50 TTTCTTGTG ATCAGCACCA TCACTACCAC TGCTCTCTTC AAAGCCACCA CGTCTGTCC 1080
 CAGGATGGTT GCAACAACCA CCATAGGGAC TTTTGTGCTC TACTTCCACA CAATAGCCAG 1140
 AGTAAGCTTT TGAATATGTA GGTCAAGTCA TGTCTCTCTC TTCTCTTCAA AACCTCCGA 1200
 TGGCTTTTCA TATTACTCAA AAGAAAACCT AAAACTTTGC TGTGAGATCT ATGTGACCCG 1260
 GCTTATTCTT CCTCTACTT TATCTCTGTA TTGCTCTTCC TCACTCTACT CCAGCCATCC 1320
 55 CACTCTCTTG CTGCTTGTC TATACTCCTA AAAGAAGTTC AGTCTTCCCT TATGATATTT 1380
 TCCAGGTTAG AAGTCATGGA GCTGGATCT AAATCCATGT CAGTCTGACT ATGAGTTCTG 1440
 CACCTTCTTA TTCAACCCCA TTGCTTAGAG GTGCTTGATT GCTCAATAAT AGATTCCATG 1500
 GACACAGTCA GCTCTTCTG AGAAAAGGCA GCTCAGCAT TCCATGAGAT CCGCACATCC 1560
 60 TTTTGAGAA GAAAAAC

Seq ID No: 185 Protein sequence:
 Protein Accession #: none found

65 1 11 21 31 41 51
 VELVSMCLBY TKPINIPSKK VLSLKS

Seq ID NO: 186 DNA sequence
 Nucleic Acid Accession #: NM_002203.2
 Coding sequence: 43-3588 (underlined sequences correspond to start and stop codons)

75 1 11 21 31 41 51
 CTGCAACCC AGCGCAACTA CGGTCCCCCG GTCAGACCCA GGAATGGGGCC AGAACGGACA 60
 GGGGCCGCGC CGCTGCCGCT GCTGCTGGTG TTAGCGCTCA GTCAAGGCAT TTTAAATTGT 120

	TTGTTGGCCT	ACAATGTTGG	TCTCCAGAA	GCAAAAATAT	TTTCCGTC	TTCAAGTGAA	180
	CAGTTTGGGT	ATGCACTGCA	GCAGTTTATA	AATCCAAAG	GCAACTGGTT	ACTGGTTGGT	240
	TCACCTCGGA	GTGCTTTCC	TGAGAACCGA	ATGGGAGATG	TGTATAAATG	TCCTGTTGAC	300
5	CTATCCCATG	CCACATGTGA	AAACTAAAT	TTGCAACTT	CAACAGCAT	TCCAAATGTT	360
	ACTGAGATGA	AAACCAACAT	GAGCCTCGGC	TTGATCTCTA	CCAGGAACAT	GGGAACCTGA	420
	GGTTTCTCTA	CATGTGGTCC	TCTGTGGGCA	CAGCAATGTG	GGAAATCAGTA	TTACACAACG	480
	GGTGTGTGTT	CTGACATCAG	TCCTGATTTT	CAGCTCTCAG	CCAGCTTCTC	ACCTGCAACT	540
	CAGCCTGCTC	CTCCCTCAT	AGATGTTGTG	GTGTGTGTG	ATGAATCAAA	TAGTATTAT	600
10	CCTTGGGATG	CAGTAAAGAA	TTTTTTGGAA	AAATTTGTAC	AAGGCTTGA	TATAGGCCCC	660
	ACAAAGACAC	AGGTGGGGTT	AATTCAGTAT	GCCAAATATC	CAAGAGTTGT	GTTTAACTTG	720
	AACACATATA	AAACCAAGA	AGAAATGATT	GTAACACAT	CCCAGACATC	CCAATATGGT	780
	GGGGACCTCA	CAACACATT	CGGAGCAATT	CAATATGCA	GAAAATATGC	CTATTCAAGC	840
	GCTTCTGTGT	GGGACGAAAG	TGCTACGAAA	GTAATGGTAG	TTGTAAGTGA	CGGTGAATCA	900
	CATGATGTTT	CAATGTTGAA	AGCTGTGATT	GATCAATGCA	ACCATGACAA	TATACCTGAGG	960
15	TTTGGCATAG	CAGTCTTGG	GTACTTAAAC	AGAAACGCC	TTGATACTAA	AAATTTAATA	1020
	AAAGAAATAA	AAGCATGCG	TAGTATTCCA	ACAGAAAGAT	ACTTTTCTCA	TGTGCTGAT	1080
	GAAGCAGTCT	TACTAGAAAA	GGCTGGGACA	TTAGGAGAAC	AAATTTTCAG	CATTGAAGGT	1140
	ACTGTTCTAG	GAGGAGACAA	CTTTCAGATG	GAAATGTAC	AAGTGGGATT	CAGTGAGAT	1200
	TACTCTCTCT	AAAATGATAT	TCTGATGCTG	GGTGCACTGG	GAGCTTTTGG	CTGAGGCTGG	1260
20	ACCATGTTCC	AGAAAGATC	TCATGGCCAT	TTGATCTTTC	CTAAACAAGC	CTTTGACCAA	1320
	ATTCTGACAG	ACAGAAATCA	CAGTTCATAT	TTAGGTTACT	CTGTGGCTGC	AAATTTCTAT	1380
	GGAGAAAGCA	CTCACTTTGT	TGCTGTGCTC	CCTCGGGCAA	ATTATACCGG	CCAGATAGTG	1440
	CTATATAGTG	TGAATGAGAA	TGGCAATATC	ACGGTATTC	AGGCTCACCG	AGGTGACCA	1500
25	ATTGGCTCTC	ATTTTGGTAG	TGTGCTGTGT	TCAGTTGATG	TGGATAAAGA	CACCTTACCA	1560
	GACGTGCTCT	TGTTAGGTGC	ACCAATGTAC	ATGAGTGACC	TAAAGAAAGA	GGAGGAGAGA	1620
	GTCTACCTGT	TTACTATCAA	AAAGGGCATT	TTGGGTGAGC	ACCAATTTCT	TGAAGGCCCC	1680
	GAGGGCATTG	AAAACATCG	ATTTGGTTCA	GCAATTCAG	CTCTTTCAGA	CATCAACATG	1740
	GATGGCTTTA	ATGATGTGAT	TGTTGGTTCA	CCACTAGAAA	ATCAGAAATC	TGGAGCTGTA	1800
30	TACATTTACA	ATGCTCATCA	GGGCACTATC	CGCACAAAGT	ATTCCAGGAA	AATCTTGGGA	1860
	TCCTGCTGAG	CCCTTAGGAG	CCATCTCCAG	TACTTTGGGA	GGTCTTGGGA	TGGCTATGGA	1920
	GATTTAAATG	GGGATTTCCAT	CACCGATGTC	TCTATTGGTG	CCTTTGGACA	AGTGTTCATA	1980
	CTCTGGGTGC	AAAGATTGTC	TGATGTAGCT	ATAGAAGCTT	CATTACACCC	AGAAAAATCT	2040
	ACTTTGGTCA	ACAAGATGTC	TCAGATAAAT	CTCAACTCT	GCTTCACTGC	AAAGTTCCAGA	2100
35	CCTATTAAGC	AAAACATCA	AGTGGCCATT	GTATATAACA	TCACACTTGA	TGCAGATGGA	2160
	TTTTCATCCA	GAGTAACCTC	CAGGGGGTTA	TTTAAAGAAA	ACRATGAAAG	GTGCTGCGAG	2220
	AAGAATATGG	TAGTAAATCA	AGCACAGAGT	TGCCCCGAGC	ACRATCAATTA	TATACAGGAG	2280
	CCCTCTGATG	TTGTCAACTC	TTTGGATTGG	CGTGGGACA	TCAGTCTGGA	AAACCTGGGC	2340
	ACTAGCCCTG	CCCTGGAAGC	CTATTCGAG	ACTGCCAAGG	TCTTCAGTAT	TCCTTTCCAC	2400
40	AAAGCCTTGT	GTGAGGATGG	ACTTTGCATT	TCTGATCTAG	TCTTAGATGT	COGACAAATA	2460
	CCAGCTGTCT	AGAGACAACC	CTTTATTTGTC	AGCAACCAAA	ACAAAAGGTT	AACATTTTCA	2520
	GTAACTGAGT	AAATAAAGAG	GGAAAGTGCA	TACAACACTG	GAATGTGTTG	TGATTTTTC	2580
	GAAAGCTTGT	TTTTTGATCT	ATTTCTCCCTA	CGGTGTGATG	GGACAGAAGT	AACATGCCAG	2640
	GTGGCTGTCT	CTCAGAGTGC	TGTTGCCTGC	GATGTAGGCT	ACCCTGCTTT	AAAGAGAGAA	2700
45	CAACAGGTGA	CTTTTACTAT	TAACTTTGAC	TTCAATCTTC	AAAACCTTCA	GAATCAGGCG	2760
	TCTCTCATTT	TCCAAGCCTT	AAGTGAAGGC	CAAGAAGAAA	ACAAAGCTGA	TAATTTGGTC	2820
	AACCTCAAAA	TTCTCTCTCT	GTATGATGCT	GAAATTCAT	TAACAAGATC	TACCAACATA	2880
	AATTTTATAG	AAATCTCTTC	GGATGGGAAT	GTTCCTTCA	TCTGTGACAG	TTTTGAAGAT	2940
	GTGGTCTGCA	AAATCACTCT	CTCCCTGAAG	GTAACAACAG	GAAGTGTCTC	AGTAAGCATG	3000
50	GCAACTGTAA	TCATCCACAT	CCCTCAGTAT	ACCAAGAGAA	AGAACCACCT	GATGTACCTA	3060
	AAACCTGAGT	AAACAGACAA	GGCTGGTGAC	ATCAGTTGTA	ATGCAGATAT	CAATCCACATG	3120
	AAATAGGAC	AAACATCTTC	TTCTGTATCT	TTCAAAAGTG	AAAATTTTCAG	GCACACCAAA	3180
	GAATTTGAAT	GCAGAACTGC	TTCTGTGATG	AATGTTACCT	GCTGTTGAA	AGACGTTTAC	3240
	ATGAAGGAGG	AATACTTTGT	TAATGTGACT	ACCAGAAAT	GGAAACGGAC	TTTCGCATCA	3300
55	TCACCTGTTG	AGACAGTACA	GCTAACGGCA	GCTGCAGAAA	TCAACACCTA	TAAACCTGAG	3360
	ATATATGTGA	TTGAAGATAA	CACTGTTTACG	ATTCCCTTGA	TGATATGAA	ACCTGATGAG	3420
	AAAGCCGAGG	TACCAACAGG	AGTTATAATA	GGAAGTATAA	TTGCTGGAAT	CCCTTTGCTG	3480
	TTAGCTCTGG	TTGCAATTTT	ATGGAAGCTC	GGCTCTCTCA	AAAGAAATA	TGAAAGATG	3540
	ACCAAAAATC	CAGATGAGAT	TGATGAGAC	ACAGAGCTCA	GTAGCTGAC	CAGCAGACCT	3600
60	ACCTGCAGTG	GGAAACGGCA	GCATCCGAGC	CAGGGTTTGC	TGTTTGCCTG	CATGGAATTC	3660
	TTTTTAAATC	CCATATTTT	TTTATCATGT	CGTAGGTAAA	CTAACCTGGT	ATTTTAAAGAG	3720
	AAAACCTGAG	GTCAGTTTGG	ATGAAGAAAT	TGTGGGGGT	GGGGGAGGTG	CGGGGGGAG	3780
	GTAGGAAAA	AATAGGAAA	ATACCTATTT	TATATGATGG	GGGAAAAAAA	GTAATCTTTA	3840
	AACTGGCTGG	CCCAGAGTTT	ACATTCTAAT	TTGCATTGTC	TCAGAAACAT	GAAATGCTTC	3900
65	CAAGCATGAC	AATTTTAA	GAAAAATATG	ATACTCTCAG	ATTTTAAAGG	GAAAACTGT	3960
	TCTCTTTAAA	ATATTTGTCT	TTAAACAGCA	ACTACAGAG	TGGAAGTGCT	TGATATGTAA	4020
	GTACTTCCAC	TTGTGTATAT	TTTAAATGAT	ATTGATGTTA	ACAGAGGGGG	AAAACAAAAC	4080
	ACAGGTTT	TCAATTTATG	CTGCTCATCC	AAAGTTGCCA	CAGATGATAC	TTCCAAAGTA	4140
	TAATTTTATT	TATAAATAG	GTAAATTTTG	TGTTGGTTC	CTTTTATACC	ACGGCTGCC	4200
70	CTTCCACACC	CCATCTTGCT	CTAATGATCA	AAACATGCTT	GAATTAAGTGA	GCTTAGAGTA	4260
	TACCTCTTAT	ATGTCATTT	AAATTTAGGAG	AGGGGGGAT	ATAGAGACTA	AGGCACAAA	4320
	TTTTTGTAA	AATCAAGAT	ATAACATTTA	TGTAAATCC	CATCTGCTAG	AGGCCATCC	4380
	TGAGGAGGAG	GAGGAAATTT	CCTTTCTCTT	TTAGGAGGCA	CAACAGTTCT	GATGAGTAA	4440
	CTTCTAGGAT	TTGTTGGCT	GACTGGCAGT	AACTAGTGA	ATTTTGAAG	GATGAGTAA	4500
75	TTCTTTGGCA	ACCTCTCTCC	TCCCTTACTG	AACTACTCTC	CCACTCTCTG	GTGGTACCAT	4560
	TATTTATAGAA	GCCCTCTACA	GCCTGACTTT	CTCTCCAGCG	GTCCAAAGTT	ATCCCTCTCT	4620
	TTACCCCTCA	TCCAAAGTTC	CCACTCCTTC	AGGACAGCTG	CTGTGATTA	GATATTAGGG	4680

5 GGGAAAGTCA TCTGTTTAAT TTACACACTT GCATGAATTA CTGTATATAA ACTCCTTAAC 4740
 TTCAGGGAGC TATTTTCATT TAGTGCTAAA CAAGTAAGAA AAAAAGCTA GAGTGAATTT 4800
 CTAATGTTTG GAATGTTATG GATGTAAAC AATGTAAAGT AAAACACTCT CAGGATTTC 4860
 CCAGAAAGTTA CAGATGAGGC ACTGGAAACC ACCACCAAT TAGCAGGTGC ACCTTCTGTG 4920
 GCTGTCTTGT TTCTGAAGTA CTTTCTCTTC CACAAGAGTG AATTGACCT AGGCAAGTTT 4980
 GTTCAAAAGG TAGATCTCGA GATGATTGG TCAGATTGGG ATAAGGCCCA GCAATCTGCA 5040
 TTTTAACAG CACCCAGTC ACTAGGATGC AGATGGACCA CACTTTGAGA AACACCAACC 5100
 ATTTCTACTT TTGCACTT ATTTCTCTG TTCTGAGCC CCCACATTCT CTAGGAGAAA 5160
 10 CTTAGATTAA AATTCACAGA CACTACATAT CTAAGCTTT GACAAGTCT TGACCTCTAT 5220
 AAACCTCAGA GTCTCATTA TAAATGGGA AGACTGAGCT GAGTTTCAG AGTGTATGCT 5280
 TTTAGTTTGA AAGTCTATG ATCTGATCTG GACTTCTAT AATACAAATA CACAATCTCT 5340
 CAAGAATTTG ACTTGGAAAA G

Seq ID NO: 187 Protein sequence:

Protein Accession #: NP_002194.1

1 11 21 31 41 51
 20 MGPERTGAAP LPILLVLALS QGILNCCLAY NVGLPEAKIP SGPSSEQFGY AVQQFINPKG 60
 NWLLVSGPWS GFPPNRMGDV YKCPVDLSTA TCEKLAQLTS TSIPNVTEMK TMSLGLLIT 120
 RNMGTGGFLT CGPLWAQCGG NQYTTGVCS DISDFQLSA SPSPATQPCP SLIDVVVCD 180
 ESNISYPMDA VKNPLEKPVQ GLDIGPTKQ VGLIQYANNP RVVFNLTYYK TKEKIVATS 240
 25 QTSQYGGDLT MTFGAIQYAR KYAYSAAAGS RRSATKVMVU VTDGESHGDS MLKAVIDQCN 300
 HDNLRPFQIA VLGYLNRNAL DTQNLKEIK AIASIPTRY PPNVSDAAL LEKAGTLGEG 360
 IFSIBGTVOG GDNFQMEMSQ VGFSADYSSQ NDILMLGAVG AFGNSGTIVQ KTHSHLIFP 420
 KQAFDQILQD RHHSSYLGS VAAISTGEST HFVAGAPRAN YTGQVLYSV NENGNITVIQ 480
 AHRGDDIGSY FGSLVCSVDV DKDTITDVLV VGAPMYMSDL KKEGRLVYLF TIKKILGQH 540
 QLEGGPELIE NTRPGSAIAA LSDINMDGPN DVIUGSPLEN QNSGAVIYIN GHQGTIRTKY 600
 30 SQKLGSQDGA FRSHLYPFGR SLOGYDLNG DSITDVSIGA FGQVQLWSQ SIADVALRAS 660
 FTPEKITTUN KNAQILKLC FSAKFRPTKQ NNQVAIVYNI TLDADGFSR VTSRGLFKEN 720
 NERCLQKMMV VHQQSCPEH IYIQBPSPDV VNSLDLRVDI SLENPGTSPA LEAYSETAKV 780
 FSIPTFKKCCG EAGLCISDLV LDVRQIPAAQ EQPFIVSNQN KRLTFSTLTK NKRESAYNTG 840
 35 IVDVDFENLF FASPSLEVDG TEVTCQVAAQ QKSVACDVGY PALKREQQVT FTINFDNLQ 900
 NLQNSQSLF QALSESQEN KADNLVNLKI PLLYDABEHL TRSTNINPFE ISSDGNVPSI 960
 VHSFEDVDFK FIFSLKVTG SVPVSMATVI IHIPQYTKK NPLMYLTVGV TDKAGDLSK 1020
 ADINPASKIQ TSSSVSPKSE NFRHTKELAC RTASCNSVTC WLKDVHMKGR YFVNVTIRW 1080
 NGTFASSTFQ TVQLTAAAEI NTYNPSIYVI EDNTVTIPLM IMKPEKAEV PTGVILGSII 1140
 AGILLLALLV AILNKLGFKK RYKRMKMP DEIDETTEL S

Seq ID NO: 188 DNA sequence

Nucleic Acid Accession #: NM_002210.1

Coding sequence: 42-3188 (underlined sequences correspond to start and stop codons)

45 1 11 21 31 41 51
 GGCTACCGCT CCCGCTTGG CGTCCGCGC GCACCTCGGC GATGGCTTTT CGGCCGGGC 60
 GACGGCTGCG CTCTGGTCCC CGCGGCTCC CGCTTCTTCT CTGGGACTCT CTGCTACCTC 120
 TGTGCGCGCG CTTCACACTA GACGTGACA GTCTGCGGA GTACTCTGGC CCCGAGGAA 180
 50 GTTACTTCCG CTTCGCGGTG GATTCTCTCG TGCCGAGCGC GTCTTCCCGG ATGTTTCTTC 240
 TCGTGGGAGC TCCCAAGCA AACACACCC AGCCTGGGAT TGTGGAAGGA GGGCAGGTCC 300
 TCAATGTGTA CTGTCTTCT ACCCGCGGT GCCAGCCAAT TGAATTGAT GCAACAGGCA 360
 ATAGAGATTA TGCCAAGGAT GATCCATTGG AATTAAAGT CCATCAGTGG TTTGGAGCAT 420
 CTGTGAGGTC GAAACAGGAT AAAATTTTGG CCGTGCCCC ATTTGACCAT TGGAGAACCTG 480
 55 AGATGAAACA GGAGCGAGAG CCGTTGGAA CATGCTTTCT TCAAGATGGA ACAAGACTG 540
 TTGAGTATGC TCCATGTAGA TCACAAGATA TTGATGCTGA TGGACAGGGA TTTTGTCAAG 600
 GAGGATTGAG CATGTATTTT ACTAAGCTG ACAGAGTACT TCTTGGTGGT CCGGTAGCT 660
 TTTATTTGCA AGGTGAGCTT ATTCGAGTC AAGTGCAGA AATCGTATCT AAATACGACC 720
 CCAATGTTTA CAGCATCAAG TATAATAACC AATTAGCAAC TCGGACTGCA CAACTATTT 780
 60 TTGATGACAG CTATTTGGGT TATCTGTGG CTGTGAGGA TTTCAATGGT GATGCTATAG 840
 ATGACTTTGT TTCAAGAGTT CCAAGAGCAG CAAGGACTTT GGGATGGTT TATATTTAG 900
 ATGGGAAGAA CATGTCTCC TTATACAATT TACTGGCGA GCAGATGGCT GCATATTTCC 960
 GATTTCTCTG AGCTGCCACT GACATTAATG GAGATGATTA TGCAGATGTG TTTATTTGAG 1020
 CACCTCTCTT CATGGATCGT GGCTCTGATG GCAACTCCTA AGAGGTGGGG CAGGTCTCAG 1080
 65 TGTCTCTACA GAGAGCTTCA GGAGACTTCC AGAGACAAA GCTGAATGGA TTTGAGTCT 1140
 TTGACCGGTT TGGCAGTGCC ATAGCTCCTT TGGGAGATCT GGACCAGGAT GGTTCATAG 1200
 ATATTGCAAT TGCTGCTCCA TATGGGGGTG AAGATAAAA AGGAATTTGT TATATCTTCA 1260
 ATGGAAGATC AACAGGCTTG AACGCACTCC CATCTCAAAT CCTTGAAGGG CAGTGGGCTG 1320
 CTGGAAGCAT GCCACCAAGC TTGGCTATT CAATGAAAGG AGCCACAGAT ATAGACAAA 1380
 70 ATGATATATC AGACTTAATT GTAGGAGCTT TTGGTGTAGA TCGAGCTATC TATACAGGG 1440
 CCAAGACAGT TATCACTGTA AATGCTGCTC TTGAAGTGTG CCCTAGCATT TTAATCAAG 1500
 ACAATAAAC CTGCTCACTG CTGGAACAG CTCTCAAGT TTCTGTGTTT AATGTTAGGT 1560
 TCTGCTTAAA GGCAGATGGC AAGGAGTAC TTCCAGGAA ACTTAATTTC CAGGTGGAAC 1620
 TGTCTTTTGA TAACTCAAG CAAAGGAGG CAATTCAGCG AGCACTGTTT CTCTACAGCA 1680
 75 GGCCCCCAGT TCACTCCAAG AACATGACTA TTTCAAGGG GGAAGTATG CAGTGTAGG 1740
 AATTGATAGC GTATCTGCGG GATGAATCTG AATTAGAGA CAAACTCACT CCAATTACTA 1800
 TTTTATAGGA ATATCGGTTG GATTATAGAA CAGCTGCTGA TACAACAGGC TTGCAACCCA 1860

TTCTTAACCA GTTCACGCGT GCTAACATTA GTCCACAGGC TCACATTCTA CTTGACTGTG 1920
 GTGAAGACAA TGCTGTGAAA CCCAAGCTGG AAGTTCTGT AGATAGTGT CAAAGAAGA 1980
 TCTATATTGG GGATGACAA CTTCTGACAT TGAATTGTAA GGCTCAGAAT CAAGGAGAAG 2040
 GTGCCTACGA AGCTGAGCTC ATCGTTTCCA TTCCACTGCA GGCTGATTTC ATCGGGGTTG 2100
 5 TCCGAAACAA TGAAGCCTTA GCAAGACTTT CTTGTGCATT TAAGACAGAA AACCAAACTC 2160
 GCCAGGTGGT ATGTGACCTT GGAACCCAA TGAAGGCTGG AACTCAACTC TTAGCTGGTC 2220
 TTGTTTCAG TGTCGACCGG CAGTCAGAGA TGGATCTTC TGTAATATT GACTTACAAA 2280
 TCCAAAGCTC AATCTATTTT GACAAAGTAA GCCCAGTTGT ATCTCACAAA GTTGATCTTG 2340
 CTGTTTTCAG TGCACTTGAG ATAAGAGGAG TCTCGAGTCC TGATCATATC TTCTTCCGAA 2400
 10 TTCCAAACTG GGAGCACAGG GAGAACCCTG AGACTGAAGA AGATGTTGGG CCAGTTGTTT 2460
 AGCACATCTA TGAGCTGAGA AACAATGGTC CAAGTTTCAIT CAGCAGGCA ATGCTCCATC 2520
 TTCAGTGGCC TTACAAATAT AATAATACCA CTCCTTGTGA TATCTTTCAT TATGATATTG 2580
 ATGACCAAT GAACCTGCACT TCAGATATGG AGATCAACCC TTTGAGAATT AAGATCTCAT 2640
 CTTTGCAAC AACTGAAAAG AATGACCGG TTGCGGGCA AGGTGAGCGG GACCATCTCA 2700
 15 CTCTTAAGCG GATCTTGCC CTCACTGAAG GAGATATICA CACTTTGGGT TGTGAGTTG 2760
 CTCAGTGCTT GAAGATTGTC TGCCAAGTTG GGAGATTAGA CAGAGGAAG AGTGCATCTT 2820
 TGATCACTTA GTCATTACTG TGGACTGAGA CTTTTATGAA TAAAGAAAAT CAGATTCATT 2880
 CCTATTCTCT GAAGTGTCTT GCTTCATTTA ATGTCATAGA GTTTCCTTAT AAGATCTTTC 2940
 20 TATGAGAGA TATCACCAAC TCCACATGG TTAOCATTA TGTCACTGG GCAATTCAGC 3000
 CAGCGGCCAT GCCTGTGCTT GTGTGGTGA TCAATTTAGC AGTTCAGCA GATGTTTGC 3060
 TACTGGCTGT TTTGGTATT GTAAATGACA GATGGGCTT TTTTAAACGG GTCCGGCCAC 3120
 CTCAGAGAAGA ACAGAAAGG GAGCAGCTTC AACCTCATGA AAATGGTGAA GGAACCTCAG 3180
 AAACCTTACT GCACTTTTAA AGTTATGCTA CATCTTGACC CACTAGAATT AGCAACTTTA 3240
 25 TTATGATATT AAACCTTCTT CATGAGGAGT AAAAATCCAA GGCTTTACTG CTGATAGTGC 3300
 TAATTGGCAT TAACCAACAA ATAGAGATTA TATTTGTCAA CCTTCTCCTT ATAAATAAGT 3360
 TCAGACATAC ATTTAATAAC ATAGGCTGAC TTGTGTTTT AGGTATTAA ATATAAAT 3420
 TTCAGGGAT AGTTTTTATT CAATGTATAT AAGACAGGTA GTGCCTGATT TACTACTTTA 3480
 TATAAATAG TACCTCCTTC AGTTACTGTT TCTGATTTAA TGTAAGGAAC TTTATTTGTT 3540
 30 GTGTGTTGT TGTGTTGTT TGTGTTTAA AAGCAGTCCA AATTGGGACC TTAGCAATCA 3600
 TAGTCTTTGT ATAGGTACTT AATGTTAATA CATATACAC TACAGTTTAC TTTTCAGAA 3660
 ATCAAGAGCT TTATAACTGC ATGAACCTGG ATTTTITTA TCACTCATAT GGTAGAATTT 3720
 TATAACACA TACATGATAC CATCCAAAT CTGCTTTTA ATACAAAGG TACAATATT 3780
 TGTTTTATGA TGAATACTG TGAATCCTA TTCACTTCT GTTTATATTA AATCCCAAT 3840
 35 ATTTTATTAC ATTTTAACT TGTATAAAT TTAGGTCAA TCCCTCAAGC CAACCTATAC 3900
 TAAAAATAG TTCCATAATC ACAAATGCT CTTTGTGTA ATTTGTTAAT TTCACCTGAA 3960
 TATCATAATG CTTAAAGCCA TATGAGTTG GAAATATTAT CCAAGCATA TTTATTCAT 4020
 TGTTTTAGTC TGGCTATTTA CAGTATAAAA AAAGCATTTT ATTAATAATC TGTGTAGTTC 4080
 40 TTTGAGATAG TTGCTATGTC ATATAGTAG TATTACATTC TTAGAGTAGA GCAGAGTTT 4140
 TAGTTAGTAT TAAATTTATT TCTCCATTC ATGTACTTTT CCTTATATT CCAAACTGT 4200
 TACTGAGAT GGGTCAAGAT CAGTGAGAAA TCTTTACAGT TGACAGGAAC CTGAGCCCT 4260
 TACCCCAACT TTATAGTAA TGCTTGAAT AAAAAGCTCT TAAGGCAACT CACTGATTTA 4320
 CTTCTAGCAA TAGCATGATG TTACAGGAAT ATTAACCTCT TTTAAGCAAG GTAATGTGTA 4380
 45 AAATCAGTCT CGGCTGTCAG AATAACTTCT AAAAGGTATT TTTATAAGCA GTTCAAGTTA 4440
 CTGAAAACCT TTAAACCTT TCTGAAGTTC GTTAGTATAA ATTACTTTTC TAGGATTACT 4500
 AATAAAAGCC ACATAGGTG CAAGTTGTAG TTTTATATGG CTCTGTAGAG TGGTGAACCT 4560
 TCTAGAGGAA TATATGATT ATTCAGATT CCTCAAGGCC TGGGATGAT GATCAGTTAT 4620
 50 CCTCTTTTGT GTGCAATTAC ATCATGTTGT ACATTAGAAA TGGAGAGTTT AATAGCTCTT 4680
 TAACGTCTGT CCTCATTAGG TAATGATAAA TATTTCCCTT AAATAATTGA CTTATTTGCT 4740
 GTGTTTAAAG AATGATTGAA ATTTATCTTG CCATATCTCA TAATTTCATG CACAAGTTGA 4800
 CTGAGCTAAT CTTGAGAAAT TATTCGTAA ATAGGAGCAC ATTTAGTTGA GGTATACAG 4860
 GTAGGACTCT AGACAAAACC TTCTATTTTA GCTTTAGTGA ATTTCAAAG TAATGGTCT 4920
 TGGAGTATAG ATTTTATTA GTAGCTTGAA AGAGCTTAAT CATATGCAAT AAGTATTTT 4980
 55 ATTACCAATA AATTTAAAT TTTTAAAGAA AAATATTTT ATCTTAGGGC CAAGTTTGC 5040
 TGCCACCAA TCAGTAAGTT AGTCTATAAC AAATTTTACC CTAACAGTTT TACCACCTAG 5100
 CACAGTCAT TTCTGAAAT ATGTTGGATA GAAAGTCACT CTTTGGCAAA AGTUTTAGAA 5160
 TTTGCTTTTG TGCCATCTAT TCCTTTATG GCATCTATCT TGAAGTAAT CTTGATTGG 5220
 AGATTGAAAG ATGCTGTAT TTAGAAATTA ACATGATATC TTAATTAACC TTTATGAAT 5280
 60 ATAGTTTGT ATATAGCAT AGATTTTCTC TCAAAAATG AACATTATA TATCTACAAA 5340
 AATATGAGA AGAGCAATTT GAAAGCCTAC TTTCTGAAGA AAATGGTGGG ATTTTITTT 5400
 ATCATGATTA AATATCAAAA AATTGCCCTA TGAACACTTT AAATCTCTAA AACATTGAA 5460
 ATACTACCAT ATTTGTGATT TATGAGAAT AAAATCCAT TTTGAATGT AAAATTTTA 5520
 TGATCTGATT CAGTTTAAAG AAAACATGAA TGAACATGAA GATATTAAA ACATTGACA 5580
 65 TTGCTAAGAA ATATGATAC TGATATTGAT TTTTATATAG GTATTATT CAGATTGAT 5640
 ATTTTGAGAA AATATCATGT GAGTCATTTT TTCTGTTCTT CTTTCTCTT AACATTATC 5700
 ACTGTAATTC TGAATCT

Seq ID NO: 189 Protein sequence:
 Protein Accession #: NP_002201.1

70 1 11 21 31 41 51
 | | | | |
 MAPPPRRRLR LGPRGLPLLL SGLLLPLCRA FNLVDSPAE YSGPEGSYFG FAVDFVFPSPA 60
 SSRNLLVGA PKANTTPQPI VEGQVQLKCD WSSTRRCQPI EPDATGNRDY AKDDPLEFES 120
 HQWFGASVRS KQDKILACAP LYHWRTMKQ EREPVGTGFL QDGTKTVEYA PCRSQDIDAD 180
 75 GQFQCQGGFS IDPTKADRLV LGGPGSFYQ GQLISQVAB IVSKYDPNVY SIKYNNGLAT 240
 RTAQAI PDDS YLGSVAVG D PNDGIDDFV SGVPRRAARTL GMVYIYDGN MSSLYNFTGE 300

QMAAYPGFSV AATDINGDDY ADVFIGAPLF MDRGSDGKIQ EVGQVSVSLQ RASGDFQTTK 360
 LAGFEVFPARF GSAIAPLGLD DQDGFNDIAI AAPYGGEDKK GIVYIFNGRS TGLNAVPSQI 420
 LRGQWAARSM PPSFGYSMKG ATDIDKNGYP DLIVGAPGVD RAILYRARFV ITVNAGLEVY 480
 5 PSILNQDNKT CSLPGTALKV SCFNVRFCLK ADGKGVLPKX LNFQVELLLD KLKQKGAIRR 540
 ALFLYSRSPS HSKNMTISRG GLMQCEELIA YLRDESEPRD KLTPTITIME YRLDYRTAAD 600
 TTGLQPIILQ FTPANISRQA HILLDOGEDN VCKFKLEVSV DSDQKIIYIG DDNPLTLIVK 660
 AQMQGEGAYE ASLIVSIPLQ ADFIGVVRNN BALARLSCAF KTNQTRQVV CDLGNFMKAG 720
 TQLLAGLEFS VHQSSEMDTS VKFDLIQSS NLFDKVSPVV SHKVDLAVLA AVSIEGVSSP 780
 DHIFLPIPNW EHKNPFTSR DVGVVQVHIY ELKNGPSSP SKAMHLQWP YKYNDFLLY 840
 10 ILHYDIDGPM NCTSDMEINP LRIKISLQT TEKNDTVAGQ GERDHLITKR DLALSEGDIH 900
 TLCCGVAQCL KIVCQVGRLD RQKSAILYVK SLLWTRTFPM KENQHSYSYL KSSASPMVIE 960
 PPYKMLPIED ITMSTLVTTN VTWGIQAPM FVFWVWVILA VLAGLLLLAV LVFVNTYRMP 1020
 FKKVRPPQKE QEREQLQPHE MGEENSET

15 Seq ID NO: 190 DNA sequence
 Nucleic Acid Accession #: NM_004864
 Coding sequence: 26-952 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 20 CGGAACGAGG GCAACCTGCA CAGCCATGCC CGGCAAGAA CTGAGGACGG TGAATGGCTC 60
 TCAGATGCTC TGGTGTGTGC TGGTGTCTCT GTGGCTCCCG CATGGGGGGG CCTGTCTCT 120
 GGCCGAGGGG AGCCGCGCAA GTTTCGCGGG ACCCTCAGAG TTGCACTCGG AAGACTCCAG 180
 ATTCCGAGAG TTGCGGAAC GCTACGAGGA CCTGCTAACC AGCCTGCGGG CCAACCTAGG 240
 25 CTGGGAAGAT TCGAACACCG ACCTCGTCCC GGCCTCTGCA GTCCGGATAC TCACGCCAGA 300
 AGTGCAGCTG GGATCCGCGG GCCACCTGCA CTTGCGTATC TCTCGGGCGG CCTTCCCGGA 360
 GGGGCTCCCC GAGGCTCTCC GCTTCACCG GCGTCTGTTC CGGCTGTCCC CGACGGCGTC 420
 AAGGTCTGTG GAGGTGACAC GACGCTGCG GCGTCACTC AGCCTTGCAA GACCCCAAGC 480
 GCCCGCGCTG CACTGCGAC TGTCCGCGCC GCGCTGCGAG TCGGACCAAC TGCTGGCAGA 540
 30 ATCTTCGTCC GCACGGCCCC AGCTGGAGTT GCACCTGCGG CGCAAGCGG CCAGGGGGCG 600
 CGCAGAGAGG CGTCCGCGCA ACGGGGAGGA CTGTCCGCTC GGGCCCGGGG GTTGTCTCGG 660
 TCTGCACAGG GTCCGCGCGT CGCTGGAAGA CTTGGGCTGG GCGGATTGGG TGCTGTGCGC 720
 ACGGGAGGTG CAATGACCA TGTGCATCGG CGCGTGCCCG AGCCAGTTCC GGGCGGCAAA 780
 CATGCTCCAG CAGATCAAGA CGAGCTGCA CGGCTGAAG CCGGACACGG AGCCAGCGCC 840
 35 CTGCTGCTGG CCGCCAGCT ACAATCCAT GGTGCTCATT CAAAGACCG ACACCGGGGT 900
 GTCCTCCAG ACCTATGATG ACTTGTTAGC CAAAGACTGC CACTGCATAT GAGCAGTCTC 960
 GTGCTCTCCA CTGTGCACCT GCGCGGGGGA GCGGACTTCA GTTGTCTCGC CCGTGTGAAT 1020
 GGCCTCAAGG TTCTGTAGAC ACCCGATTCC TGCCCAACA GCTGTATTTA TATAAGTCTG 1080
 40 TTAATTATTA TTAATTATTA GGGGTGACCT TCTGGGGGAC TCGGGGGCTG GTCTGATGGA 1140
 ACTGATGATT TATTAAACAC TCTGCTGATA AAAATAAGC TGTCTGAAGT GTTAAAAAAA 1200
 AAAA

Seq ID NO: 191 Protein sequence
 Protein Accession #: NP_004855
 45 1 11 21 31 41 51
 MPQQLRLTVN GSQMLLVLLV LSWLPHGGAL SLARASRASP PGPSELHSED SRPRELNRKY 60
 EDLLRLRLAN QSWDSNTDL VPAPAVRIIT PEVRLGSGGH LHLRISRAAL PBOLPEASRL 120
 50 HRLALFRLSPT ASRSMDVTRP LRLQLSLARP QAPALHLRLS PPSQSDQLL AESSSARPOL 180
 ELHLRPPQAR GRRRARARNG DDCPLPGRC RLHETVRASL EDLGWADMVL SPREVQVIMC 240
 IGACPSQFRA ANMHAQIKTS LHLRKFDTSP APCCVPASYN PMVLIQXTDT GVSLQTYDDL 300
 LAKDCHCI

55 Seq ID NO: 192 DNA sequence
 Nucleic Acid Accession #: XM_061731.1
 Coding sequence: 1-567 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 60 ATGAGAAAAA GAAATGAGGG AGAGAACACA GAAGAGGGCA GGCTTGCTCA GCTTGCTCAA 60
 AGAAAGTTTC TCAAAGAAGA TGCCATTACA TTGCACATCT CTCTGTGTCT CTCTATTGCT 120
 GTAAAGAAGC CTTTCTCTCT GATTGGACTT GACACACAGA AGGATCTCAG TAAAGATTGG 180
 CTGTGTGTTA TGTCCACAGA CACTGGCAAG GACAGGTTTA CCAACATACT GCTGTACAC 240
 65 TCCCTCCCAA TGTGCACCAA ATCAGTAAAT AATGGGGATA ATGACTCCCC TGCCCTTACA 300
 TGGGGTGGCA AAGACACCAAG GAGCAATACT GATCTTCCCTA TCAGAGACCC TGGGGGCAAG 360
 AGTCTTTCAC TCACCAACA TTCCACCAAG CCGTCCCTG AGCATCAGTG TGACCAAGA 420
 GAGGTCTTCC AGCCACTTTC AGAGCCAGGT GTAGAAGCAG AGATGGAAGT GTTCGCTGAT 480
 70 GCTGGATGGT GGATTTATCA GAGCTGTGAG GTTCTTCTCT CAACCTTGC AAGAAAGAAG 540
 ATGATTATTT CTAAAGAAAC TGAGTGA

Seq ID NO: 193 Protein sequence
 Protein Accession #: XP_061731.1

75 1 11 21 31 41 51

MRKGNBENT EBRILAQLAQ RKFLKEDGIT LHSLSLCSIA VKPPFSLIGL DTQKDLKDL 60
 LLLMSDTDTGK DRPTNILLSH SPPMCTKSRK NGDNDSFAPT WGGKDTBSNT DLPIRDPGGK 120
 SLSLTKHSHK FVPEHQCDQR EVFQPLSEPG VEARMEVFAD AGWMIYQSCQ VPSSTLARKK 180
 MVYSKETE

5

Seq ID NO: 194 DNA sequence

Nucleic Acid Accession #: NM_005415.2

Coding sequence: 371-2410 (underlined sequences correspond to start and stop codons)

10 1 11 21 31 41 51
 | | | | |
 GAGCTGTCCC CGGTGCCGCC GACCCGGGCC GTGCCGTGTG CCCGTGGCTC CAGCGCGTGC 60
 CGCCTCGATC TCCTOGTCTC CGCTCCGCC CTCCCTTTTC CCTGGATGAA CTGGGTCTCT 120
 TTCTCTCTCT CGCATGGAA TTCTGCTCCG TGCTTTTAGC CCTCTGAGC CAAAGAAACC 180
 15 CCAGACAACA GATGCCATA GCGAGCGTAT AGCAGTAAT CCCAGCTCG GTTCTGTGTC 240
 CGTAGTTTAC AGTATTTAAT TTTATATAAT ATATATTATT TATTATAGCA TTTTGTATAC 300
 CTCATATTCT GTTTACACAT CTTGAAAGGC GCTCAGTAGT TCTCTTACTA AACAACTACT 360
 ACTCCAGAGA ATGGCAACGC TGATTACAG TACTACAGCT GCTACCGCCG CTCTGGTCC 420
 TTTGGTGGAC TACCTATGGA TGCTCATCTT GGGCTTCATT ATTGCATTG TCTTGGCATT 480
 20 CTCCGTGGGA GCGAATGATG TAGCAAAATC TTTTGTGTACA GCTGTGGGCT CAGGTGTAGT 540
 CAAAGCTGAAG CAGGCTGCA TCCTAGCTAG CATCTTTGAA ACAGTGGGCT CTGTCTTACT 600
 GGGGGCCAA GTGAGCGAAA CCATCCGGAA GGGCTTGATT GAGCTGGAGA TGTACAACTC 660
 GACTCAAGGG CTACTGATGG CCGGCTCACT CAGTCTATG TTTGTTCTG CTGTGTGGCA 720
 ACTCGTGGCT TCGTTTTTGA AGCTCCCTAT TTCTGGAACC CATGTATTG TGTGTCAAC 780
 25 TAATGTGTTT TCCTCTGTTG CAAAGGGGCA GGAGGTTGTC AAGTGTCTG AACTGATAAA 840
 AATTGTGATG TCTTGGTTTG TGTCCCACT GCTTTCTGGA ATTATGTCTG GAATTTTATT 900
 CTTCGTGTTT CGTGCATTCA TCCTCCATAA GGCAGATCCA GTTCTTAATG GTTTGGGAGC 960
 TTTGCCAGTT TTCTATGCGT GCACAGTTGG AATAAACCTC TTTTCCATCA TGTATCTGG 1020
 AGCACCGTTG CTGGGCTTTG ACAAACTTCC TCTGTGGGT ACCATCTCA TCTCGTGGG 1080
 30 ATGTGCACTT TTCTGTGCC TTTATCGTCTG GTTCTTTGTA TGTCCAGGA TGAAGAGAAA 1140
 AATTGAACGA GAAATAAAGT GTAGTCCCTC TGAAGAGCCC TTAATGGAAA AAAAGATAG 1200
 CTTGAAGAAA GACCATGAAG AAACAAAGTT GTCTGTGGT GATATTGAAA ACAAGCATCC 1260
 TGTTCCTGAG GTAGGGGCTG CCACTGTGCC CCTCCAGGCT GTGGTGGAGG AGAGAACAGT 1320
 35 CTTGAAGAG GAAACACGCA TAGATAGCAC CGTGAATGGT GCGATGCACT TGCCTAATGG 1380
 GAACCTGTTC CAGTTCAGTC AAGCGTCAAG CAACCAATA AACTCCAGTG GCCACTCCA 1440
 GTATCACACC GTGCTAAGG ATTCCGGCCT GTACAAAGAG CTACTCCATA AATTACATCT 1500
 TGSCAAGTGG GGAGATTGCA TGGGAGACTC CGGTGACAAA CCCTTAAGGC GCAATATAG 1560
 40 CTATCTCTCC TATACCATGG CAATATGTGG CATGCCCTCTG GATTCAATCC GTGCCAAGA 1620
 AGGTGAACAG AAGGGCGAAG AAATGGAGAA CGTGACATGG CCTAATGCAG ACTCCAAGAA 1680
 GCGAATTGGA ATGGACAGTT ACACCAAGTA CTGCAATGCT GTGTCTGACC TCTACTCAGC 1740
 ATCTGAGATA GACATGAGTG TCAAGGCAGC GATGGGTCTA GGTGACAGAA AAGGAAGTAA 1800
 TGGCTCTCTA GAAGAATGGT ATGACCAGGA TAAGCCTGAA GTCTCTCTCC TCTTCCAGTT 1860
 45 CCCTCAGATC CTTACAGCCT GCTTTGGGTC ATTCCGCCAT GGTGGCAATG ACGTAAGCAA 1920
 TGCATTTGGG CTTCTGGTGT CTTTATATTT GGTTTATGAC ACAGGAGATG TTTCTTCAA 2040
 AGTGGCAACA CCAATATGGC TTCTACTCTA TGGTGGTGT GGTATCTGTG TGGTCTGTG 2100
 GGTTTGGGGA AGAAGAGTTA TCCAGACCAT GGGGAAGGAT CTGACACCGA TCACACCTC 2160
 TAGTGCTCTC AGTATTGAAC TGGCATCTGC CCTCACTGTG GTGATTGCAT CAAATATGG 2220
 50 CCTCCCATC CAGTACAAC ACATTGAAAGT GGGCTCTGTT GTGTCTGTG GCTGGCTCCG 2280
 GTCCAGGAAG GCTGTGACT GGGTCTCTT TGTGAACATT TTTATGGCCT GGTGTGCTAC 2340
 AGTCCCATCT TCTGGAGTTA TCAGTCTCTC CATCATGGCA ATCTTCAGAT ATGTCTCCT 2400
 CAGAATGTGA AGCTGTTTGA GATTAAAAAT TGTGTCAATG TTTGGGACCA TCTTAGGTAT 2460
 55 TCCTGCTCCC CTGAAGAATG ATTACAGTGT TAACAGAGA CTGACAAGAG TCTTTTATT 2520
 TGGGAGCAGA GGAGGGAAGT GTTACTTGTG CTATAACTGC TTTTGTGCTA AATATGAATT 2580
 GTCTCAAAAT TAGCTGTGTA AAATAGCCCG GGTTCACACTG GCTCCTGCTG AGGTCCCTT 2640
 TCCTTCTGGG CTGTGAATTC CTGTACATAT TTCTCTACTT TTTGTATCAG GCTTCAATTC 2700
 CATATGTTT TAATGTTGTC TCTGAAGATG ACTTGTGATT TTTTCTCTT TTTTAAATC 2760
 CATGAAGAGC CGTTTGACAG AGCATGCTCT GCGTGTGTGG TTTCCAGAG TCTGCTCTC 2820
 60 ACATGCACAG GGATTAAACA ACRAAAATAT AACTACAAC TCCCTGTAG TCTCTTATAT 2880
 AAGTAGAGTC CTTGTACTC TGCCCTCTCT TCACTAGTGG CAGGATCTAT TGGCATATTC 2940
 GGGAGCTTCT TAGAGGGATG AGGTCTCTT AACACAGTGA AAATTTAAAT TAGTAACITT 3000
 TTTGCAAGCA GTTTATTGAC TGTATTGCT AAGAAGAAAT AAGAAAGAAA AAGCCTGTG 3060
 GCAATCTTGG TTATTCTTT AAGATTCTG GCACTGTGGG ATGGATGAAT GAAGTGGAA 3120
 65 GTGAACCTTG GGCAAGTTAA ATGGGACAGC CTTCATGTT CATTTGTCTA CCTCTTAAT 3180
 GAATAAAAAA GCCTACAGTT TTTAGAAAAA ACCCGAATTC

Seq ID NO: 195 Protein sequence

Protein Accession #: NP_005406.2

70 1 11 21 31 41 51
 | | | | |
 NATLITSTTA ATAASGPLVD YLMLILGFI IAFVLAFLSVG ANDVANSFGT AVSGGVVTLK 60
 QCTLASIFE TVGSVLLGAK VSTIRKGLI DVMYNSYQG LLMAGSVSAM FGSVAVQLVA 120
 75 SPFLKPISTG HCIVGATIGF SLVAKQEGV KWSLEIKIVM SWFVPSLLSG IMSGILFPLV 180
 RAFTLHKADP VPGRLRALPV FYACTVQINL FSIYMTGAPL LGFDKLPWLG TILISVGCAV 240
 FCALIVNPFV CPRMKRKIER EIKCSFSESP LAKFKNSLKS DHEBTKLSVG DIENKHPVSE 300

5 VGPATVPLQA VVEERTVSFK LGDLERAPER ERLPSVDLKE ETSIDSTVNG AVQLPANGNLV 360
 QFSQAVSNQI NSSGHSQYHT VHKDSGLYKE LRLKHLAKV GDCMGDSGDK PLRRNNSYTS 420
 YTMACGMPL DSFRKREGEQ KGEEMKLTW PNADSKRIR MDSYTSYCNA VSDLHSASHI 480
 DMSVKAAMGL GDRKGSNGSL KENYDQDKPE VSLLPQFLQI LTACPGSPAH GGNVSNIAIG 540
 10 FLVALYLVVD TGDVSSKVT PIWLLYGGV GICVGLWVG RVVIQTMGKD LTPITPSSGF 600
 SIELASALTV VIASNIGLPI STTHCKVGSV VSVGNLRSKK AVDWRLFRNI FMAHFVTVEI 660
 SGVISAALMA IFRYVILRM

Seq ID NO: 196 DNA sequence

Nucleic Acid Accession #: NM_000020.1

Coding sequence: 283-1794 (underlined sequences correspond to start and stop codons)

15 1 11 21 31 41 51
 AGGAAACGGT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA 60
 AGAAACATT TTGCTCCAGC CCCCATCCCA GTCCCGGGAG GCTGCCGCGC CAGCTGGCGC 120
 GAGCGAGGCC CTCGCCGCT CCAGCCCGGT CCGGGCGCGC GCGGACCCCG AGCCCGCGT 180
 CCAGCGCTGG CGGTGCAACT GCGGCCGCGC GGTGGAGGGG AGGTGGCCCC GTCCCGCGA 240
 20 AGCTAGCGC CCGGCCACCC GCAGAGCGGG CCCAGAGGGA CCATGACCTT GGGCTCCCCC 300
 AGGAAAGGCC TTCTGATGCT GCTGATGGCC TTGTTGACCC AGGGAGACCC TGTGAAGCG 360
 TCTCGGGGCC CGCTGGTGAC CTGCACGTGT GAGAGCCAC ATTGCAAGGG GCCTACCTGC 420
 CGGGGGCGCT GGTGCACAGT AGTGTGTGTG GGGAGGAGG GAGGACCCCG CCAGGAACAT 480
 CGGGCTGGG GGAACCTGCA CAGGAGGCTC TGCAAGGGGC GCCCCACCGA GTTGTCTAAC 540
 25 CACTACTGCT GGCACAGCCA CCTCTGCAAC CACAACGTGT CCCTGTGCTG GAGGGCCACC 600
 CAACTCTCTT CGGAGCAGCC GGGAAACAGT GGCCAGCTGG CCCTGATCCT GGGCCCCGTG 660
 CTGGCTCTGC TGGCCCTGGT GGCCCTGGGT GTCTGGGCC TGTGGCATGT CGAGCGAGG 720
 CTGAGGAGC AGCGTGGCCT GCACAGCGAG CTGGAGAGT CCAGTCTCAT CCTGAAGCA 780
 TCTGAGCAGG GGCACAGCAT GTTGGGGGAC CTCTGGACA GTGACTGCAC CACAGGAGT 840
 GGCTCAGGGC TCCCTTCTCT GGTGCAGAGG ACAGTGGCAC GGCAGGTGTC CTGTGTGAG 900
 30 TGTGTGGGAA AAGGCCGCTA TGGCGAAGTG TGGCGGGCT TGTGGCACGG TGAGAGTGTG 960
 GCGCTCAAGA TCTTCTCTCT GAGGATGAA CAGTCTGCT TCGGGAGAC TGAGATCTAT 1020
 AACACAGTAT TGTCTAGACA CGACAACATC CTAGGCTTCA TCGCTCAGA CATGACCTCC 1080
 CCAACTCGA GCACGAGCT GTGGCTATC AGCCTATACC ACAGACCGG CTCCCTCTAC 1140
 35 GACTTCTCTC AGAGCAGAC GCTGGAGCCC CATCTGGCTC TGAGGCTAGC TGTGTCCGG 1200
 GCATGCGGCC TGGCGCACT GCACGTGGAG ATCTTCGTA CACAGGGCAA ACCAGCCATT 1260
 GCCACCGCGC ACTTCAAGAG CCGCATGTG CTGCTCAAGA GCAACCTGCA GTGTTCATC 1320
 GCGGACCTGG GCTGGCTGTG GATGCACTCA CAGGCGAGG ATTACCTGGA CATCGGCAAC 1380
 AACCGAGAGG TGGGACCAA GCGGTACATG GCACCGAGG TGCTGGACGA GCAGATCCGC 1440
 40 AGGACTGCT TTGAGTCCTA CAAGTGGACT GACATCTGG CCTTGGCCT GGTGCTGTG 1500
 GAGATTGCC CCGCGACCAT CGTGAATGGC ATCTGGAGG ACTATAGACC ACCCTTCTAT 1560
 GATGTGTGTC CCAATGACCC CAGCTTTGAG GACATGAAGA AGGTGTGTG TGTGGATCAG 1620
 GCACCCCAAC CCAATCCCTAA CCGGCTGGCT GCAGACCGG TCCTCTCAGG CTAAGTCTAG 1680
 ATGATGCGGG AGTGTGCTA CCAAAACCCC TCTGCGGAC TCACCGGCT GCGGATCAAG 1740
 45 AGACACTAC AAAAATTAG CAACAGTCCA GAGAAGCCTA AAGTGAATCA ATAGCCAGC 1800
 AGACACTGAT TCCTTTCTGC CTGAGGGGGG CTGGGGGGGT GGGGGGAGT GATGTGTGTC 1860
 CTATCTGGGT AGAGGTAGT TGAGTGTGCT GTGTGCTGGG GATGGGAGC TGGCCCTGCC 1920
 TGCTCGGCC CCAGCCACC CAGCCAAAA TACAGCTGGG CTGAACCTG

Seq ID NO: 197 Protein sequence

Protein Accession #: NP_000011.1

55 1 11 21 31 41 51
 MTLGSPRKGL LMLLMALVTQ GGPVKPSRGP LVTCTCBSPH CKGPTCRGAW CTVVLVREEG 60
 RHPQERKRG NLRRLCRGR PTEFVNHYCC DSHLCNHNVS LVLEATQPPS EQPGTDGQLA 120
 LILGFLVALL ALVALGVGLL WHVRRRQKQ RGLHSELGHS SLILKASEQG DTMLGDLIDS 180
 DCTTGGSGSL PFLVQRTVAR QVALVECVGK GRYGEVVRGL WHGESVAVKI FSSRDEQSWF 240
 60 RETSIYNTVL LRHDNILGFI ASDMTSRNSS TQLMLITHYH EHGSYVDFLQ RQTLRPHLAL 300
 RLAYSAAACGL AHLRYMIFGT QGKPAIAHRD PKSRNVLVKS NLQCCIDILQ LAVMHSQGGD 360
 YLDYGNPFRV GTRRYMAPEV LDEQIRTDCE ESYKMDIWA PGLVLNRIAR RTIVNGIVED 420
 YRPPFYDVVP NDSPEDMKK VVCVDQQTPT IPNRLADPV LSLAQMMRE CMYPNPBARL 480
 TALRIKKTILQ KISNSPEKPK VIQ

Seq ID NO: 198 DNA sequence

Nucleic Acid Accession #: NM_003199.1

Coding sequence: 200-2203 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 CCGGGGGATC TTGGCTGTGT GTCTGCGGAT CTGTAGTGGC GCGGGGCGCG GCGGCGCGG 60
 GGAGGCAGCA GCGCGCGGAG CCGGCGCAGG AGCAGGCGGC GCGGTGCGC GCGGCGGTTA 120
 GACATGAAGC CGCGCTCGGC CCGCGCGGTG CACGAGAGGC CCCTTCTGCG GCGGCGGCGG 180
 75 TTTGTGTGAT TTTGCTAAAA TGCAATACCA ACAGCGAATG GCTGCTTAG GAGGAGACAA 240
 AGAGCTGAGT GATTACTGCG ATTTCACTGC GATGTTTTC CCTCTGTGA GCAGTGGGAA 300
 AATGGACCA ACTTCTTTG CAACTGGACA TTTACTGCG TCAATGTAG AAGACAGAAG 360
 TAGCTCAGG TCCTGGGGGA ATGAGGAGCA TCCAAGCCCG TCCAGGAAC ATGGAGATGG 420

GACTCCCTAT GACCACATGA CCAGCAGGGA CCTTGGGTCA CATGACAATC TCTCTCCACC 480
 TTTTGTCAAT TCCAGAAATAC AAAGTAAAC AGAAAGGGG TCATACTCAT CTATGGGAG 540
 AGAATCAAA TACAGGGTT GCCACAGCA GAGTCTCCTT GGAGGTGACA TGGATATGGG 600
 CAACCCAGGA ACCCTTTCGC CCACCAACC TGGTCCAG TACTATCAAT ATTCTAGCAA 660
 5 TAATCCCGCA AGGAGGCCTC TTCAAGTAG TGGCATGAG GTACAGACAA AGAAAGTTCG 720
 AAAAGTTCCT CCAGGTTTGC CATCTTCAAT CTATGCTCCA TCAGCAAGCA CTGCGACTA 780
 CAATAGGGAC TGGCCAGGCT ATCTTTCCTC CAACCCAGCA ACCAGCACTT TCCCTAGCTC 840
 CTCTCTCATG CAGATAGGCC ATCAGAGCAG TGACCTCTGG AGCTCTCTCA GTGGGATGAA 900
 TCAGCCTGGC TATGCAGGAA TGTGGGCAA CTCTCTCAT ATTCCACAGT CCAGCAGCTA 960
 10 CTGTAGCCTG CATCCACATG AACGTTTGG CTATCCATCA CACTCTCAG CAGACATCAA 1020
 TTCCAGTCTT CCTCCGATGT CCACCTTCCA TCGTAGTGT ACACCACTT ACAGCACCTC 1080
 TTCTCTAGC CTCTCTGCA ACAGGACAGA CAGTATAAT GCACCAAGAG GAAGCGGGGC 1140
 AGCCGGGAGC TCCAGAGCTG GAGATGCTCT GGGGAAGCA CTGCTTGA TCTATTCTCC 1200
 AGATCAGACT AACACAGCT TTTTATCAA CCCTTCAACT CCGTTGGCT CTCTCTCATC 1260
 15 TCTCTCAGCA GGCACAGCTG TTTGGTCTAG AAATGGAGGA CAGGCTCAT CGTCTCTTAA 1320
 TTATGAAGGA CCTTACACT CTTTGCAAG CGAATTTGAA GATCGTTAG AAAGACTGGA 1380
 TGATGCTATT CATGTTCTCC GGAACCATG AGTGGGCCA TCCACAGCTA TGCGTGTGG 1440
 TCATGGGAGC ATGCAAGGA TCAATGGACC TTCTCAATAT GGAGCCATGG TGGTCTGGG 1500
 CTCAGGTTAT GGAACCGGCC TTCTTTCAGC CAACAGACAT TCATCATGG TGGGACCCA 1560
 20 TCGTAGAGAT GCGCTGGCCC TGAGAGGAGC CCATCTCTT CTGCCAAACC AGGTTCGGT 1620
 TCCAGAGCTT CCGTCCAGT CTGCGACTTC CCTGACCTG AACCCACCCC AGGACCTTTA 1680
 CAGAGCAGTG CCACCCAGGAC TACAGGGGCA GAGTGTCTCC TCTGCGAGCT CTGAGATCAA 1740
 ATCCGATGAC GAGGTTGATG AGAACCTGCA AGACACGAAA TCTTCGAGG ACAAGAAAT 1800
 AGATGACGAC AAGAAGGATA TCAATCAAT TACTAGCAAT AATGACGAT AGGACCTGAC 1860
 25 ACCAGAGGAG AAGCAGAGC GTGAGAAGGA CGGAGGAGT GCCACCAATG CCGAGAGCG 1920
 TCTCGGGTC COTGACATCA ACGAGGCTTT CAAGAGCTC GCGCGCATGG TGCGCTCCA 1980
 CCTCAAGAGT GACAGCCCC AGACCAAGCT CCGTATCTC CACCGAGCGG TGGCGCTCAT 2040
 CCTCAGTCTG GAGCAGCAAG TCCGAGAAAG GAATCTGAAT CGAAGAGCTG CGTGTCTGAA 2100
 AAGAGGGAG GAGAGAGGAG TGTCTCGGA GCCTCCCTCT CTCTCTTGG CCGGCCACA 2160
 30 CCTGGAATG GAGAGCGCAT CGAATCAGT GGGACAGAT TAAAGGGTC CAGTGTGCA 2220
 CATGCTCTCA TTAACAAGAG AGACCACTTC CTTAACAGCT GTATTATCTT AAACCCAT 2280
 AAACACTTCT CCTTAACCCC CATTTTGTG ATATAAGCA AGTCTGAGT GTTATGAAT 2340
 GCAGACGCAA GAGGTTTCAG CATTCCTCAT TATCAAAAA CAGAAAAACA AAAAAAGAA 2400
 35 AGAAAAAGT GCACTTTGAG GGACGACTTT CTTTAAATA TCATTAGAA TGTGCAAGC 2460
 AGTATGTACA GGTGAGACA CAGCCAGAG ACTGAACGGC

Seq ID NO: 199 Protein sequence:

Protein Accession #: NP_003190.1

40 1 11 21 31 41 51
 | | | | |
 MHQQRMAAL GTDKELSDLL DFSAMFSPPV SSGKNGPTSL ASGHFTGSNV EDRSSSGSWG 60
 NGGHPSPSRN YGDGTPYDHM TSDRLGSHDN LSPFPVNSRI QSKTERGSYS SYGRESNLQG 120
 CHQQSLLGGD MDMMNPGLTS PTKPGSQYQ YSSNNPFRRE LHSSAMEVQT KKVKKVPGL 180
 45 P8SVYAPAS TADYNRDSFG YPSSKPATST FPPSPFMQDG HHSDDPMSSS SGMNQPGVAG 240
 MLGNSHPTQ SSSYCSLHPH ERLSYPSHSS ADINSSLEPM STEHRSCTNH YSTSCTPPA 300
 NQDTSIMANR GSGAGSSQT GDALGKALAS TYPDHTNHS FSNPSTPVG SPPSLSAGT 360
 VMSRNGQAS SPSNYEGPLH SLQSRIEDRL ERLDDAIHVL RNHAVGPSTA MPGGHGMHG 420
 IIGPSHNGAM GGLGSGYGTG LLSANRHSIM VGTREDGVA LRGSNLSLFP QVFPQLPVQ 480
 50 SATSPDLNPP QDPYRGWFPQ LQGQSVSSGS SEIKSDDEGD ENLQDTKSSS DKKLDDDKED 540
 IKSITSNDD BDLTPBQKAE RBKRRMANN ARERLRVRED NEAFKELGRM VQLHLKSDKP 600
 QTKLLLLHQA VAVILSLEQQ VRERNLNPKA ACTKRREBEK VSSEPPPLSL AGPHPMGDGA 660
 SNHMGGM

Seq ID NO: 200 DNA sequence

Nucleic Acid Accession #: BC005987 (1-1286), BE888744 (1287-1756)

Coding sequence: 124-525 (underlined sequences correspond to start and stop codons)

60 1 11 21 31 41 51
 | | | | |
 GGCAGAGAG GAAGATTCTT GAGAGATGCA GCTGCCTGAA CCGAGCCCTG CCGAACAGCT 60
 GAGATTGCA CTGCAACCAT GAGTGAGAAC AATAAGAAAT CCTTGGAGAG CAGCTACGG 120
 CAACTAAAT GCCATTTCAC CTGGAACCTG ATGAGGGAG AAAACTCCTT GGATGATTT 180
 GAGACAAAG TATTTTACCG GACTGAGTT CAGAATGTTG AATTCAAAGC CACAATGTGC 240
 65 AACCTACTGG CCTATCTAAA GCACCTCAAA GGGCAAAAGC AGGCAGCCTT GGAATGCTTA 300
 CGTAAAGCTG AAGAGTTAAT CCAGCAAGAG CATGCTGACC AGGCAGAAAT CAGAAGCTG 360
 GTCACCTGGG GAAACTATGC CTGGGTCTAC TATCACATGG GCGACTCTC AGACGTTTCA 420
 ATTATGTAG ACAAGGTGAA ACATGCTCTT GAGAAGTTT CCACTCCCTA TAGAATTGAG 480
 AGTCCAGAGC TTGACTGTGA GGAAGGGTGG ACACGGTTAA AGTGTGGAG AAACCAAAAT 540
 70 GAAAGAGCGA AGGTGTGCTT TGAGAAGGCT CTGGAAGAGA AGCCAAAGAA CCCAGAAATC 600
 ACCTCTGGAC TGGCAATAGC AAGCTACCGT CTGGACAACT GGCCACCATC TCAGAACGCC 660
 ATTGACCCCT TGAGGCAAGC CATTCGGCTG AATCTTGACA ACCAGTACCT TAAAGTCTCT 720
 CTGGCTCTGA AGCTTCATAA GATGCTGAAA GAAGGTGAAG AGGAAGGTGA AGGAGAGGAG 780
 TTAGTTGAAG AAGCCTTGGA GAAAGCCCA GGTGTAACAG ATGTACTTCC CAGTGACAGC 840
 75 AGTGTTCATC GAAGAAAGA TGAGCCAGAC AAAGGATGAT AACTGCTTAA AAAGGCTTTA 900
 GAATACATAC CAAACAATGC CTACCTGCAT TGCCAAATG GGTCTGCTA TAGGCAAAA 960

5
10
15

```
GTCTTCCAAG TAATGAATCT AAGAGAGAAT GGAATGTATG GGAAGAAGAA GTTACTGGAA 1020
CTAATAGGAC ACCTGTGGC TCATCTGAAG AAAGCTGATG AGGCCAATGA TAATCTCTTC 1080
CGTGTCTGTT CCATCTCTGC CAGCCTCCAT GCTCTAGCAG ATCAGTATGA AGAAGCAGAG 1140
TATTACTTCC AAAAGGAATT CAGTAAAGAG CTTACTCTGT TAGCGAAACA ACTGCTCCAT 1200
CTCGGTTATG GCACCTTTCA GCTCTACCAA ATGAGTGTGT AAGACAAGGC CATCCACCAC 1260
TTTATAGAGG GTGTAAAAAT AAACCAAGAA TCAAGGAGAA AAGAAAAGAT GAAAGACAAA 1320
CTGCAAAAAA TTGCCAAAAA GCGACTTTCT AAAATGGAG CAGATTCTGA GGCTTTGCAT 1380
GTCTTGGCAT TCCTTCAGGA GCTGAATGAA AAAATGCAAC AAGCAGATGA AGACTCTGAG 1440
AGGGGTTTGG AGTCTGGAAG CCTCATCCCT TCAGCATCAA GCTGGAATGG GGAATGGAAGA 1500
ATAGAGATGT GGTGCCCACT AGGCTACTGC TGAAGGGAG CTGAATTCCT TCCCAAGATT 1560
GGTATTCAAA ATATGTAATG ACTGATATGG CAAAAGATTG GACTAAGACA CTGGCCATAC 1620
CACTGACAAA GGTATATGTT AACCTGAATT GCTGGTCTCT AAAAGAGCCC AAGGAATCTCT 1680
GGGAGAGGGA CAGATTGGGG GGTCTGCCAG GGCTGCGCTA AATTATTCTC AATGATTGTG 1740
CTCTTGGCGG AACTTC
```

Seq ID NO: 201 Protein sequence:

Protein Accession #: AAA59191

20
25
30

```
1 11 21 31 41 51
| | | | |
MSRMNKNLE SSRLQLKCHF TWNLMBGENS LDDPDKVFF RTEPQNRBFK ATMCNLLAYL 60
KHLKQNEAA LECLEKAEEL IQQEHADQAE IRLVTVNGVY AMVYYHMGRL SDVQIYVDKV 120
KHVCEKSSFP YRISPELDC BEGWRLKCG GNQNRKAVC PEKALEKPKF NPEPTSLAI 180
ASYRLDNWPP SQNAIDPLRQ AIRLNPNDQY LKVLALALKH KMREGEGBEG EGKLVSEAL 240
EKAPGVTVL RSAAKFYRK DBPKAIELL KKALEYIPNN AYLCQIGCC YRAKVPQVMN 300
LRBNMGVGR KLELIGHAV AHLKADKAN DNLFRVCSIL ASLHALADQY SDASYFPKE 360
FSKELTPVAK QLEHLRYGNF QLYQMKCEDK AIHFFIEGVK INQKSREKEK MKDKLQKIAK 420
MRLXNGADS EALHVLAFLO ELNEKMQQAD EDSERGLESG SLIPSASSWN GE
```

Seq ID NO: 202 DNA sequence:

Nucleic Acid Accession #: NM_003090

Coding sequence: 57-824 (underlined sequences correspond to start and stop codons)

35
40
45
50
55

```
1 11 21 31 41 51
| | | | |
GAATTCGCGG GGAGGCCACG GGCTTTCCAC AGCGCGGGGG AACGGGAGGC TGCAGGATGG 60
TCAGCTGACG GCGGAGCTGT ATCGAGCAGG CGGCGCAGTA CACCAACGCG GTGCGCGACC 120
GGAGCTTGA CCTCGGGGG TATAAAATTC CCGTCATTGA AAATCTAGGT GCTACGTTAG 180
ACCACTTTGA TGCTATTGAT TTTTCTGACA ATGAGATCAG GAAACTGGAT GGTTTTCCTT 240
TGTGAGAAG ACTGAAAACA TTGTTAGTGA ACAACAACAG AATATGCCGT ATAGGTGAGG 300
GACTTGATCA GGCTCTGCCC TGCTGACAG AACTCATCTC CACCAATAAT AGTCTCGTGG 360
AACTGGTGR KCTGGACCTT CTGGCATCTC TCAAATCGCT GACTTACCTA AGTATCTTAA 420
GAAATCCGCT AACCAATAAG AAGCATTACA GATTGTATGT GATTATATAA GTTCCGCAAG 480
TCAGAGTACT GGATTTCCAG AAAGTGAAAC TAAAGAGCG TCAGGAAGCA GAGAAAATGT 540
TCAGGGCAA ACGGGGTGCA CAGCTTGCAA AGGATATGCG CAGGAGAAGC AAAACTTTTA 600
ATCAGGTGCG TGCTTTGCCA ACTGACAAA AGAGAGGTGG CCGATCTCCA GGGATGTAG 660
AAGCAATCAA GAATGCCATA GCRAATGCTT CAATCTGCGC TGAAGTGGAG AGGCTGAAGG 720
GGTGTCTGCA GTCTGTGAG ATCCCTGGCA GAGAACGCGC ATCAGGGCCC ACTGATGATG 780
GTGAGAAGA GATGGAAGAA GACACAGTCA CAACCGGTC CTGAGCAGTG AGGCAGATGT 840
ATAATAATG GCCCTCTTGG AACAACTCTT GCTTTTGGAA CATGGTATAA TAGCCTTGTT 900
TGTGTAGCA AAGTGGAAATC TATCAGCATT GTTGAAATGC TTAAGACTGC TGCTGATAAT 960
TTTGTATAT AAGTTTTGAA ATCTAAATGT CAATTTCTTA CAATATTATAA AATAAACTC 1020
CACTTCTAT GCTAAAAAAA AAAAAAGGA ATTC
```

Seq ID NO: 203 Protein sequence:

Protein Accession #: NP_003081.1

60
65

```
1 11 21 31 41 51
| | | | |
MKVLTAELE QAAQYTNVAV DRELDLRGYK IPVIEHLGAT LDQPDADIFS DNEIRKIDGF 60
PLLRLRLTLL VNNRIRICIQ EQLDQALPCL TELILTNLSL VELQDLPLA SLKSLTYLSI 120
LRNPVTNKKH YRLYVIYKVF QVRVLDQKV KLKERQEARL MFKGKRGQAL AKDIARRSKT 180
FNPAGLPTD KRGGPSPGD VEAIKNAIAN ASTLAEVERL KGLLQSGQIP GRERRSGPTD 240
DGEEMEREDT VTNGS
```

Seq ID NO: 204 DNA sequence:

Nucleic Acid Accession #: NM_017643.1

Coding sequence: 169-1401 (underlined sequences correspond to start and stop codons)

70
75

```
1 11 21 31 41 51
| | | | |
AATAGCAATA GCTTTATAGC AGCTCCGGTT ACCTGTTTAA AACATGGAAG GAGAGTCGCT 60
CCGAGATAGC CCTCACAGGT GCGCCCTGGAG CAGGAGATGG TGGAGCAGAT CTTCCTTGTT 120
TGGAGGAGC CTGAGGTGGA CCTCGCCTCC TGAATCTGGA AGGCACCTAT GGGGACCTCG 180
TGGGGTGATA TCTCAGAAAA TGTGAGAGTA GAAGTTCCCA ATACAGACTG CAGCCTACCT 240
```

5 ACCAAAGTCT TCTGGATTGC TGGAAATTGA AAATTAGCAG GTTACAATGC CCTTTTAAGA 300
 TATGAAGGAT TTGAAATGA CTCTGGTCTG GACTTCTGGT GCAATATATG TGGTTCTGAT 360
 ATCCATCCAG TTGGTTGGTG TGCAGCCAGC GGAAACCTC TTGTCTCTCC TAGAACTATT 420
 CAGCATTAAT ATACAAACAG GAAAGCTTTT CTAGTGAAAC GACTTACTGG TGCCAAAACA 480
 CTGCTCTCTG ATTTCTCCCA AAAGGTTTCA GAGAGTATGC AGTATCTCTT CAACCTTGC 540
 ATGAGAGTAG AAGTGGTTGA CAAGAGGCAT TTGTGTGAA CACGAGTAGC AGTGGTGGAA 600
 AGTGAATTG GAGGAAGATT AAGACTAGTG TATGAAGAAA GCGAAGTAG AACAGATGAC 660
 TTCTGGTGCC ATATGCACAG CCCATTAAATA CATCATATTG GTTGGTCTCG AAGCATAGGT 720
 CATGATCTCA AAAGATCTGA TATTACAAAG AAACAGGATG GACATTTTGA TACACCACCA 780
 10 CATTATTATT CTAAGGTAAA AGAAGTAGAC CAGAGTGGGG AATGGTTCAA GGAAGGAATG 840
 AAATTGGAAG CTATAGACCC ATTAAATCTT TCTCAATAT GTGTGCAAC CATTAGAAAG 900
 GTGTAGTGT ACGGATTCCT GATGATTGGG ATCGATGGCT CAGAAGCAGC AGACGGATCT 960
 GACTGGTTCT GTTACCATGC AACCTCTCCT TCTATTTTCC CTGTGGTTTT CTGTGAAATT 1020
 AACATGATTG AACTTACTCC ACCCCAGAGT TACACAAAAC TTCTTTTAA ATGGTTTGAC 1080
 15 TACCTCAGGG AAACCTGGCT CATTCGAGCA CCACTAAAAC TATTTAATAA GGATGTTCCA 1140
 AATCAGGAT TTCTGTAGG AATGAAATTA GAAGCAGTAG ATCTCATGGA GCCACGTTTA 1200
 ATATGTGTAG CCACAGTAAC TCGAATTATT CATCGTCTCT TGAGGATACA TTTTGATGGA 1260
 TGGGAAGAG AGTATGATCA GTGGGTAGAC GTGAGTCAC CTGACCTCTA TCTGTAGGG 1320
 TGTGTCACT TAACGTGATA TCAACTACAG CCTCCAGCAT CACAGTGTAA GTTGGTATAC 1380
 20 AGAAAGGTG TCCTTTGTGA AAATTCAGCA ATTCTCCAGA GCACTATCTC ACATAAGTCA 1440
 TCTTATGAGC TCACAGGACA AGAATATACC TATGTCGTAT GTGTTGCCAG GTAAGACATT 1500
 AAGACTCAAC AACATATCA CAGAATCAGA CCATGTGTCC CATGCCAATG TGAATCCAAT 1560
 AGTCAATTAC ATATGACTA TAGAAACACA ACAGTCACCA AATTAAACTA GACTTACTAT 1620
 25 TTTAGTAGGT TAAATATTAC ATACTAAAAG TTTATTGGTA GGTAAATAAT CCTTTAGAT 1680
 AAATAGTGA AAATGTCTCA TGTGTAGGCT ATGGTTTGT AGGAACAAGT ACCCTTATT 1740
 TCAGAGCATC ATGTACTTAA GTATAATGGT CTTGGTAAAG ATAGTTTATA TAAGTTGAT 1800
 CTAGACAACT GTATGCTCTA AATTGTAAAC AATTATCTAG TACCAATTTT CCTTTTAT 1860
 TTTTCAGCAT CAAGAGAAA CCAATCAGCT TCATCAAAAC AGAAGAAAAA GCTAAGTCC 1920
 CACCAATACA AAGACATAA GAAAGTGGG TCACCAAGTG GTGTTACAT ACATTTTCTA 1980
 30 ATGTGTTAAT AATTGGAGTC ACAGTATTCT TGGACAGAAA ATGATATATC TTGTGAGAAC 2040
 TGATGATTGT GCATTATGTA TTATGCTTAA AGGTGCAGTA TGCCATAAAA GGCAACCT 2100
 TGCAATATG AGAAACACTG ATATTTTACT AACAGAGGAA ATGATTACCA CAGTATTTAA 2160
 AGTATAGCTG GTAAGGAATA GAGTCTGTGA ATGATTCTTG AAATAATATG TAAACCTAC 2220
 TGAAGTTAA TCCTTTAA AAATTTTATT TAAAGAGAAA AATTAGCAGC CAGGTGCACT 2280
 35 GGCTCAAGCC TGTAAATCCA GCACCTTAGG AGGCCGAGGC TGGCAGATCA CAGGTGAGG 2340
 AGATCGAGAC CATCTGGCT AACACGGTGA AACCTGTCT CCACCAAAAA TACAAAAAAT 2400
 CTGCGGCGG CTGTGGCACA CGCCTGAAGT CCAGCTACT CAGGAGGCTG AGGCAAGAG 2460
 ATCATTGAA CCCAGGAGG AGAGGTTGCA GTGGGCCAAG ATCAGGCCAC TACATTCCAG 2520
 CTGGCAACA CAGCAAGACT CTGTCTCAA AAAAAAAAAA AAAA

Seq ID NO: 205 Protein sequence
 Protein Accession #: NP_060113.1

45 1 11 21 31 41 51
 MGTGMDISE NVREVPMTD CSLPTKVFHI AGIVKLAGYN ALLRYBGFEN DSGLDPWCNI 60
 CGSDIHPVGM CAAGKPLVP PRTIQHKYTN WKAFLVKRLT GAKTLFPDPS QKVSEMQYP 120
 FKPCMRVEV DKKHLCTRVR AVVESVIGGR LRLVYEESED RTDDFWCHMH SPLIHHIGHS 180
 50 RSIGHRFKRS DITRKQDGHF DTPPHLPAKV KEVDQSGSEW KBGMKLEAID PLNLSTICVA 240
 TIRKVLADFG LMIGIDGSEA ADGSDWFCYH ATSPSIFPVG PCRNIMIELT PPRGYTKLFP 300
 KWFYLYRZET SIAAPVKLPN KDVPNHGFRV GMLKLEAVDM EPRLICVATV TRIIHRLLRI 360
 HFDGVEEEDY QMWDCSPDL YFVGWQCLTG YQLQPPASQC KLVYRKGVLL

55 Seq ID NO: 206 DNA sequence
 Nucleic Acid Accession #: NM_012334
 Coding sequence: 223-6399 (underlined sequences correspond to start and stop codons)

60 1 11 21 31 41 51
 GAGACAAAGG CTGCCGTGGG GACGGGCGAG TTAGGGACTT GGGTTTGGGC GAACAAAAGG 60
 TGAGAAGGAC AAGAAGGGAC CGGGCGATGG CAGCAGGGGA GCCCGCGGG CGCGGTCCT 120
 CGGAGATGGC GCGGTGACAC GCATGGTTTC CCGGACCGCG CGCGGCGCT GACTTCGGG 180
 65 AGTCGAGCG GCACTCGGGC AGTCCGGGAC TGGCTGGAA CAATGGATAA CTTCCTCACC 240
 GAGGGAACAC GGGTCTGGCT GAGAGAAAAT GGCCAGCATT TTCCAAGTAC TGTAAATTC 300
 TGTGAGAAAG GCATCGTGGT CTTCGGGACA GACTATGGTC AGGTATTAC TTACAAGCAG 360
 AGCACAAATTA CCACCCAGAA GGTGACTGCT ATGCACCCCA CGAACGAGGA GGGCGTGGAT 420
 70 GACATGGGCT CCTTGACAGA GCTCCATGGC GGCTCCATCA TGTATACTT ATTCCAGCGG 480
 TATAAGAGAA ATCAATATA TACCTACATC GGCTCCATCC TGGCTCTCGT GAACCCCTAC 540
 CAGCCCATCG CCGGCTGTGA CGAGCCTGCC ACCATGGAGC AGTACAGCG CGGCCACCTG 600
 GGCGAGCTCG CCCGACATC CTTCGCCATC GCCAAGAGT GCTACCGCTG CCGTGTGGAAG 660
 CGCTACGACA ACCAGTGCAT CCTCATCAGT GGTGAAAGTG GGCAGGTAA AACCGAAGC 720
 75 ACTAAATTGA TCCTCAAGTT TCTGTCAATC ATCAGTCAAC AGTCTTTGGA ATTTGCTTAA 780
 AAGAGAGAGA CATCTGTGT TGAACGAGCT ATTTCTGAAA GCAGCCCAT CATGGAAGCT 840
 TTCGCGAATG CGAAGACCGT GTACAACAAC AACTCTAGT GCTTTGGGAA GTTTGTTAG 900

	CTGAACATCT	GTGAGAAAGG	AAATATTGAG	GGCGGGAGAA	TTGTAGATTA	TTTATTAGAA	960
	AAAAACCGAG	TAGTAAAGCA	AAATCCCGGG	GAAAGGAATT	ATCACATATT	TTATGCACTG	1020
	CTGGCAGGGC	TGGAACTAGA	AGAAAGAGAA	GAATTTTATT	TATCTACGCC	AGAAACTAC	1080
	CACACTGTGA	ATCAGTCTGG	ATGTGTAGAA	GACAAGACAA	TCAGTGACCA	GGAATCCCTT	1140
5	AGGGAAGTTA	TTACGGCAAT	GGACGTGATG	CAGTTCAGCA	AGGAGGAAGT	TCGGGAAGTG	1200
	TCGAGGCTGC	TTGCTGGTAT	ACTGCATCTT	GGGAACATAG	AAATTTATCAC	TGCTGGTGGG	1260
	GCACAGGTTT	CCTTCAAAAC	AGCTTTGGGC	AGATCTGGGG	AGTTACTTGG	GCTGGACCCA	1320
	ACACAGCTCA	CAGATGCTTT	GACCCAGAGA	TCATGTGTCC	TCAGGGGAGA	AGAGATCCTC	1380
	ACGCTCTCTA	ATGTTCAACA	GGCAGTAGAC	AGCAGGGACT	CCCTGGCCAT	GGCTCTGTAT	1440
10	CGCTGTGCT	TTGAGTGGGT	AATCAAGAAG	ATCAACAGCA	GGATCAAGGG	CAATGAGGAC	1500
	TTCAAGTCTA	TTGGCATCCT	CGACATCTTT	GGATTTGAAA	ACTTTGAGGT	TAATCACTTT	1560
	GAACAGTTCA	ATATAAATA	TGCAAAACGAG	AACTTCAGG	AGTACTTCAA	CAAGCATATT	1620
	TTTTCTTTAG	AACAACCTAGA	ATATAGCCGG	GAAGGATTAG	TGTGGGAAGA	TATTGACTGG	1680
	ATAGCAATAG	GAGAAATGCT	GGACTTGATT	GAGAAGAAAC	TTGGCTCCT	AGCCCTTATC	1740
15	TTTCTTTTAT	AACAATTTCC	TCAGCCACA	GACAGCACCT	TATTGGAGAA	GCTACACAGT	1800
	CAGCAATGCA	ATAACCACTT	TTATGTGAAG	CCAGAGTTG	CAGTTAAGCA	TTTGGAGATG	1860
	AAGCACTATG	CTGGAGAGGT	GCAATATGAT	GTCCGAGGTA	TCTTGGAGAA	GAACAGAGAT	1920
	ACATTTGAG	ATGACCTTCT	CAATTTGCTA	AGAGAAAGCC	GATTTGACTT	TATCTACGAT	1980
20	CTTTTGTGAC	ATGTTTCAAG	CCGCAACCAAC	CAGGATACCT	TGAATGTGG	AAGCAAACTAT	2040
	CGCGGGCTTA	CAGTCAGCTC	ACAGTTCAAG	GACTCACTGC	ATTCTTAAAT	GGCAACGCTA	2100
	AGCTCCTCTA	ATCCTTTCTT	TGTTGCTGCT	ATCAAGCCAA	ACATGCGAGAA	GATGCCAGAC	2160
	CAGTTTGAAC	AGGCGGTTGT	GCTGAACCCAG	CTGCGGTACT	CAGGGATGCT	GGAGACTGTG	2220
	AGAATTCGCA	AAGCTGGGTA	TGCGGTCCGA	AGACCTTTTC	AGGACTTTTA	CAAAAGGTAT	2280
25	ATAGTGTCTA	TGAGGAATCT	GGCTCTGCTT	GAGGAGCTGG	GAGGGAAGTG	CACGAGCCTG	2340
	CTGCACTCTT	ATGATGCTTC	CAACAGCGAG	TGGCAGCTGG	GGAAGACCAA	GGCTTTTCTT	2400
	CGAGAACTCT	TGGAACAGAA	ACTGGAGAAG	CGGAGGGAAG	AGGAAGTGAG	CCACGCGGCC	2460
	ATGGTGATTC	GGGCCCATGT	CTTGGGCTTC	TTAGCAAGAA	AACAATACAG	AAAGGTCTCT	2520
	TATTTGTGCG	TGATAATACA	GAAGAATTAC	AGAGCATTC	TTCTGAGGAG	GAGATTTTGT	2580
30	CACCTGAAAA	AGGCAAGCAT	AGTTTTCAG	AAGCACTCA	GAGGTGAGAT	TGCTCGGAGA	2640
	GTTTACAGAC	AATTTGCTGG	AGAGAAAAGG	GAGCAAGAG	AAAGAGAGAA	ACAGGAAGAG	2700
	GAAGAAGAGA	AGAAACGGGA	GGAAAGAGAA	AGAGAAAGAG	AGAGAGAGCG	AGAGGAAGCC	2760
	GAGCTCGCG	CCGACGAGGA	AGAAGAAACG	AGGAAGCAGC	AAGAAGTGA	AGCTTTCAG	2820
	AGAAGCCAGA	AGGAAGCTGA	ACTGACCCGT	GAACTGAGAA	AACAGAAAGAA	AAATAGCAG	2880
35	GTGAGAGAGA	TGCTCCGTCT	GGAGAAAGAA	ATCGAGGACC	TGCAGCGCAT	GAAGGAGCAG	2940
	CAGGAGCTGT	CGCTGACCGA	GGCTTCCCTG	CAGAAGCTGC	AGGAGCGCG	GGACAGGAG	3000
	CTCCGACGCG	TGAGGAGAGA	AGCCTGCAAG	CGCGCCGAGG	AGTTCTCTGA	GTCCCTCAAT	3060
	TTTCGAGAGA	TCGACGAGTG	TGTCGGAAT	ATCGAGCGGT	CCCTGTGCGT	GGGAAGCGAA	3120
40	TTTTCGAGAG	AGCTGGCTGA	GAGCGCATGC	GAGGAGAGGC	CCAATTTCAA	CTTCAGCCAG	3180
	CCCTACCCAG	AGGAGGAGGT	CGATGAGGGC	TTGGAAGCGC	ACGACGAGCG	CTTCAGGAC	3240
	TCCCCCAACC	CCAGCGAGCA	CGGCCACTCA	GACCAGCGAA	CAAGTGGCAT	CCGACCCAGC	3300
	GATGACTCTT	CAGAGGAGGA	CCCATACATG	AACGACAGCG	TGGTGCCAC	CAGCCCCAGT	3360
	GGGAGAGCA	CGTGCTGCT	CGCCCCATCA	GTGCAGGACT	CCGGGAGCCT	ACACAATCC	3420
	TCCCAAGCGG	AGTCCACCTA	CTGCATGCC	CAGAACGCTG	GGGACTTGCC	CTCCCCAGAC	3480
45	GGCGACTAG	ACTACGACCA	GGATGACTAT	GAGGAGCGTG	CCATCACTTC	CGGCAGCAGC	3540
	GCTGCTTCT	CCAACTCCTA	CGGCAGCCAG	TGCTCCCGCG	ACTACCGCTG	CTCTGTGGGG	3600
	ACCTCAACCA	GCTCGGGTGC	CTACCGGTTT	AGCTCTGAGG	GGCGCAGTTC	CTCGTTTGAA	3660
	GATAGTGAAG	AGGACTTTGA	TTCCAGGTTT	GATACAGATG	ATGAGCTTTT	ATACCGCGCT	3720
	GACTCTGTGT	ACAGCTGTGT	CACCTCTGCG	TATTTCCACA	GCTTTCTGTA	CATGAAGAGT	3780
50	GGCTGTATGA	ACTCTTGGAA	ACGCGGCTGG	TGCTCTCTCA	AGGATGAAG	CTTCTGTGG	3840
	TTCCGCTCCA	AGCAGGAGGC	CCTCAAGCAA	GGCTGGCTCC	ACAAAGAAAG	GGGGGGCTCC	3900
	TCCACGCTGT	CCAGGAGAAA	TTGGAAGAA	CGCTGGTTTG	TCCTCCGCCA	GTCCAGCTG	3960
	ATGTACTTTG	AAAACGACAG	CGAGGAGAA	CTCAAGGGCA	CGTAGAAGT	GGCAACGGCA	4020
	AAAGAGATCA	TAGATAACAC	CACCAAGGAG	AATGGGATCG	ACATCATTTAT	GGCCGATAGG	4080
55	ACTTTCCACC	TGATTGACAG	GTCCCCAGAA	GATGCCAGCC	AGTGGTTTCA	CGTGCTGAGT	4140
	CAGGTCCAGC	CGTCCACGGA	CCAGGAGATC	CAGGAGATGC	ATGATGAGCA	GGCAAAACCA	4200
	CAGAATGCTG	TGGGCACTTT	GGATGTGGGG	CTGATTGATT	CTGTGTGTGC	CTCTGACAGC	4260
	CCTGATAGAC	CCAATCGTT	TGTGATCATC	ACGCGCAACC	GGGTGTGCA	CTGCAACGCC	4320
	GACACGCGCG	AGGAGATGCA	CCACTGGATA	ACCCTGCTGC	AGAGGTCCAA	AGGGGACACC	4380
60	AGAGTGGAGG	GCCAGGAATT	CATCCTGAGA	GGATGTTTGC	ACAAAGAGGT	GAAGAACAGT	4440
	CGGAAGATGT	CTTCACTGAA	ACTGAAGAAA	CGGTGGTTTG	TACTCACCCA	CAATTCCTCTG	4500
	GATTACTACA	AGAGTTTCA	GAAGAAGCG	CTCAAACTGG	GGACCTTGGT	CCTCAACAGC	4560
	CTCTGCTCTG	TGCTCCCGCC	AGATGAGAA	ATATTCAAAG	AGACAGGCTA	CTGGAACGTC	4620
	ACCGTGTACG	GGGCAAGCA	CTGTTACCGG	CTCTACACCA	AGCTGTCTCA	CGAGGCCACC	4680
65	CGGTGTGACA	GTCCCATTTA	AAAGCTGACT	GACACCAAGG	CCCGATCGA	CACCCACC	4740
	CAGCAGCTGA	TTCAAGATAT	CAAGGAGAAC	TGCTGAACT	CGATGTGTT	GGACAGATT	4800
	TACCAAGCGA	ACCGGATCCT	TGATACACCC	CATCACCTCT	TGCTCTCCC	GCTCTGCCC	4860
	CTTCCGTATG	GGGACATAAA	TCTCAACTTG	CTCAAGACA	AAGGCTATAC	CACCCCTCAG	4920
	GATGAGGCCA	TCAAGATATT	CAATTCCTCTG	CAGCAACTGG	AGTCCATGTC	TGACCAATTT	4980
70	CCAATTAATCC	AGGGATCCTT	ACAGACAGGG	CATGACCTGC	GACCTCTGCG	GGAGGAGCTG	5040
	TACTGCCAGC	TTATCAAAAC	GACCAACAAA	GTGCCCCACC	CGGCAAGTGT	GGGCAACCTG	5100
	TACAGCTGGC	AGATCTGAC	ATGCTTGAGC	TGCACCTTCC	TGCCAGTTCG	AGGGATTCTC	5160
	AAATATCTCA	AGTTCCATCT	GAAGAAGATA	CGGGAACAGT	TTCCAGGAAC	CGAGATGGAA	5220
	AAATACGCTC	TCTTCACTTA	CGAATCTCTT	AAAGAAACCA	AATGCCGAGA	GTTTGTGCTC	5280
75	TCCCGAGATG	AAATAGAAGC	TCTGATCCAC	AGGCAGGAAA	TGACATCCAC	GGCTTATATG	5340
	CTGCGCGGCG	GCTCTGCAA	GATCACCATC	AACTCCCAAC	CCACTGCTGG	GGAGGTGGTG	5400
	GAGAAGCTGA	TCCAGGGCCT	GGCCATGGAG	GACAGCAGGA	ACATGTTTGC	TTTGTGTGAA	5460

5 TACAACGGCC ACGTCGACAA AGCCATTGAA AGTCGAACCG TCGTAGCTGA TGTCTTAGCC 5520
 AAGTTTGAAG AGCTGGCTGC CACATCCGAG GTTGGGGACC TGCCATGGAA ATTCTACTTC 5580
 AAACTTTACT GCTTCTCTGA CACAGACAAC GTGCCAAAAG ACAGTGTGGA GTTTGCAATT 5640
 ATGTTTGAAC AGGCCACGAA AGCGGTTATC CATGGCCACC ATCCAGCCCC GGAAGAAAAC 5700
 10 CTCCAGGCTC TTGCTGCCCT GCGACTCCAG TATCTGCAGG GGGATTATAC TCTGCACGCT 5760
 GCCATCCAC CTCTCGAAGA GGTTTATTCC CTGCAGAGAC TCAAGGCCCG CATCAGCCAG 5820
 TCAACCAAAA CCTTCACCCC TTGTGAACGG CTGGAGAAGA GCGGACGAG CTTCCTAGAG 5880
 GGGACCCCTGA GCGGAGAGCT CCGGACAGGA TCGTGGTCC GGCAGAAGGT CGAGGAGGAG 5940
 CAGATGCTGG ACATGTGGAT TAAGGAAGAA GTCTCCTCTG CTGAGCCAG TATCATTTAC 6000
 15 AAGTGGAGGA AATTTACGGG AATGAACCAG GAACAGGCCA TGGCCAAGTA CATGCCCTTG 6060
 ATCAAGGAGT GGCCTGGCTA TGGCTCGACG CTGTTTGTATG TGGAGTCAA GGAAGGTGGC 6120
 TTCCCTCAGG AACTCTGCTT GGGCTGCAGC GCGGACGCCG TCTCCGTCTA CAAGCGTGA 6180
 GAGGGAAGAC CACTGGAAGT CTTCAGTAT GAACACATCC TCTCTTTGG GGCACCCCTG 6240
 GCGAATAGCT ATAAGATCGT GGTCTGAGAG AGGAGAGCTC TCTTTGAAAC CAGTGAGGTG 6300
 20 TTGGATGTGG CCAAGCTCAT GAAAGCCTAC ATCAGCATGA TCGTGAAGAA GCGCTACAGC 6360
 ACCAGACGCT CCGCCAGCAG CAGGGGAGC TCCAGTGAA GCGGGAACAG AGCCCACTG 6420
 TCTTTGCTCT CTGAACGCAC CACCCCTGCG CTAGGCTGG CTCCAGTGG CCATGCCAG 6480
 CCAAAACAAA CACAGAGCTG CCGAGGCTTT CTGAAGCTT CTGGTCTGAG GAGGTGTCT 6540
 25 CCGAGGATCC TTTCCTCTGC CCGCTTCATT GATCCTGTAT TAAGCTGTCA ACTTTAACAG 6600
 TCTGCACAGT TTCCAAAGCT TTAAGTCTCT TAGAGGACAC ATGCCCTTAA AAAGAGGGG 6660
 AGGAACCAAG CTGCCACCAA AGCAGCCGGA AGTGCCCTAA CTGTGGAAC CAACACTAAT 6720
 CGACGTAAC TGTCTACTG AAGGGAACCT CCTTTCCTCC TTCTGGGGA GACTTAACAG 6780
 AGCGTGGAAG GGGGCGATTC TCTGTCAATG ATGCACCTAC CTCCCAACCT GATTTCCTCG 6840
 30 AATCTGAGGG AAGGTGAGGG AGTGGGAAGG GGGATGAGA GCTGAGGGG ACAGTGTGTT 6900
 TGAGCTGGAG TGTCTGCGGC AGGCTTTCTC ATGGAATGAC ATGAATCAAC TTTTCTCTT 6960
 GTTCTACTCT TTAAGTGTAC GTGCTGCTCT GTTCGTGAT GTTTCATAA ACTCAACACT 7020
 TTAATCATGG TTTCATGAGC ATTAAGAGC AAAGGGAAGA AGGATGTGTA ATGGTGTACA 7080
 CAGTCTGAT ATTTTAATAA TGCAGAGCTA TAGTCTCAAT TGTACTTTA TAAGGTGTT 7140
 35 TTATTAACAA ACCCAAATCC TGGATTTTCC TGTCTTGTCT GTATTTGAA AAACAAGTGT 7200
 TGACTCCATT GTTTTACATG TAGCAAGTCT TGCCATCTGT GTCTGCTGTA TTATAACAG 7260
 ATAAAGAGCC TACAAGATAA CTGTATTAT TAACCACTCT TCAACAGCTG GCTCCAGTGC 7320
 TGTGTTTAGA ACAAGAATGA AGTCATTITG GAGTCTTTCA TGTCTAAAG ATTTAAGTTA 7380
 AAAACAAAGT GTTACTTGA AGGTAGCTT CTATCATCTT GGATAGATTA CAGATATAAT 7440
 40 AACCATGTTG ACTATGGGG AGAGAGCTG CATTCAGAA AGCTCTTAAC ACTTGAAGTA 7500
 ATCTTCAAG GACCCGACA TTAAGTCTG AGGCTTTAAT ACACACATAT TTTATCCAA 7560
 GTTTATAATG GTGGTCTGAA CAAGGCACCT GTAAATAAAT CAGCATTTAT GACCAGAAGA 7620
 AAAATAATG GTCTTGGAC TTTTATTTT TATATGAAA AGTTTAAAG ACTTGGCCA 7680
 ACTAAGTCTA CCCACAGAA AAAAGAAAT TGCTTGTCC CTGTGTGAT ACCCATGCA 7740
 AACTGTTTGT TGCTCAGAG AAGTCTGAC AATAAAGAT ACTAGCT

Seq ID NO: 207 Protein sequence:
 Protein Accession #: NP_036466

45 1 11 21 31 41 51
 MDNFFTESTR VMLRENGQHP PSTVNSCAEG IUVFRTDYG VFTYKQSTIT HQKVTAHMT 60
 NEEGVDDMAS LTELHGGSIM YNLFORYKRN QIYTYIGSIL ASVNPYQPIA GLYEPATMEQ 120
 YSRRLHGLBP PHIFAIANEC VRCLWIKRYDN QCILISGREG AGKTESTKLI LKFLSVISQ 180
 50 SLRLSLKKKT SCVERAILES SPIMEAFGNA KTVYNNSSR FGKPVQLNIC QKGNIQGGRI 240
 VDYLLEKREY VRQNPQERNY HIFVALLAQL EHEEREFFYL STPEVHYILV QSGCVDEKTI 300
 SDQESFREVI TAMDVNQPSK EBEVREVSRL AGILHLGNI E PITAGCAQVS FKTALGRSAR 360
 LLGLDPTQLT DALTRQSMFL RGEELTFLN VQAVDSRDS LAMALYACCF EEWIKKINSR 420
 IKGNEDPKSI GILDIFGFEN FEVNHPEQFN INYANEKLQE YFNKHIFSL E QLEYSREGLV 480
 55 WEDIDWIDNG ECLDLIEKKL GLLALINKES HFPQATDSTL LEKLHSQHAN NHFVVKPRVA 540
 VNNFVGKHYA GEVQYDVRGI LEKNRDTFRD DLNLLRESR PDFIYDLFEH VSSRNQDTL 600
 KCGSKHRRPT VSSQPKDSLH SLMATLSSSN PFFVRCIKFN MQKMPDQPDQ AVVLNQLRYS 660
 GMLETVRIRK AGYAVRRPFQ DFYKRYKVLN RNLALPEVDR GKCTSLQLLY DASNSBWLQ 720
 KTKVFLRESL EQKLEKRRRE EVSHAAMVIR AHVGLFLARK QYRKVLYCVV IIQKNYRAFL 780
 60 LRRRFLHLKK AAVFQKQLR QGIARRVYRQ LLAEKREQE KKKQSEBEK KRESEERERE 840
 RRRREBELRA QQSEETRKQ ELBALQKSQK EBLTRELK E KENKQVEI LRLSKIEDL 900
 QRMKQQLSL LBSALQKLQ ERRDQLRLR EEEACRAQE FLESINPDEI DECVNIERS 960
 LSVGSEFSES LAESACEEKP NFNFSQPYPE BEVDGFEAD DDAFKDSNPF SEHGHSQDRT 1020
 65 SGIRTSDDSS REDPYNDTV VPTSPADST VLLAFSPQDS GLSHNSSGE STYCMQPNAG 1080
 DLSPDQDQD VDQDDYEDGA ITSQSSVTF S NSYGSQNSPD YRCSVGTYS SGAYRPSSE 1140
 AQSSPESSES DFDSPDTDD BLSYRRDSVY SCVTLPPYPS FLYMKGLMN SMKREKCVLK 1200
 DETFLWFRSK QEALKQWLH KKGGSSTLS RRMKKRWV LRQSKLMYFE NDSEKLGKT 1260
 VEVRTAKBII DNTTKENGID IIMADRTFHL IAESPEDASQ WFSVLSQVHA STDQBIQEMH 1320
 DEQANPNQAV GTLDVGLIDS VCASSDPDRP NSPVIITANR VLHCNADTPE EMHWITLLQ 1380
 70 RSKGDRVBS QEFIVRGWLH KEVKNSPKMS SLKLLKRWV LTHNSLDYK SSEKMLKLG 1440
 TLVNSLCSV VPPDEKIFKE TGVVNVTVYK RKHCYRLYTK LNBATRWSN AIQNVDTKA 1500
 PIDTPTQLLI QDIKENCLNS DVVEQIYKRN PILRYTHPL HSPLPLPYG DININLLKDK 1560
 GYTTLQDEAI KIFNSLQLE SMSDPIPIQ GILQTHDLR PLRDELYCQL IKQTNKVPHP 1620
 GSVGNLYSWQ LFCLSCTFL PSRGILKYLK FHLKRIREQ PGTEMKYAL FTYESLKKTK 1680
 75 CREFVPSRDE IEALHRRQEM TSTVYCHGGG SKKITINSHT TAGSEVBEKLI RGLAMEDSRN 1740
 MPALFEYNH VDKAIESRTV VADVLAKPEK LAATSEVGLD PKFYPKLYC FLDTDNVPKD 1800
 SVFPAFMEQ AHEAVIHGHH PAPERNLQVL AALRLQLQG DYTLLAAIFP LESVYSLRL 1860

KARISQSTKT FPCERLEKR RTSPLBGLR RSPRTGSVVR QKVEEQMLD MWIKEEVSSA 1920
 RASIIDKWRK PQGNMQEQAM AKYMALIKEM PGYGSTLFQV ECKRGGFPQE LNLGVSADAV 1980
 SVYKRGGRP LEVFPYEHIL SPGAPLANTY KIVVDEREILL FETSEVVDVA KLMKAYISMI 2040
 VKKRYSTTRS ASSQGSRR

5

Seq ID NO: 208 DNA sequence

Nucleic Acid Accession #: XM_059761.1

Coding sequence: 124-525 (underlined sequences correspond to start and stop codons)

10 1 11 21 31 41 51
 CGAAGATCTA TCCAAAATCA AGAAGCCTTT GATTTAGATG TTGCTGTAAA AGAAAAATAA 60
 GATGATCTCA ATCATGTGGA TTTGAATGTG TGTACAAGCT TTTCGGGCCC GGGTAGGAGT 120
 GGCATGCTCT TTAGTGAAGT TAACCTATTA AGTGGCTTTA TGCTGCTCTC AGAAGCAATT 180
 15 TCCTCTGAGCG AGACAGTGAA GAAAGTGGAA TATGATCATG GAAAACTCAA CCTCTATTTA 240
 GATTCTGTAA ATGAAACCCA GTTTTGTGTT AATATCTCTG CTGTGAGAAA CTTTAAAGTT 300
 TCAAATACCC AAGATGCTTC AGTGTCCATA GTGGATTACT ATGAGCCCAAG GAGACAGGCG 360
 GTGAGAAGTT ACAACTCTGA AGTGAAGCTG TCCTCTCTGT ACCTTTGCAG TGATGTCCAG 420
 GGCTGCCGTC CTGTGAGGGA TGGAGCTTCA GGCTCCCATC ATCACTCTTC AGTCATTTT 480
 20 ATTTCTCTGT TCAAGCTTCT GTACTTTATG GAACCTTGGC TGTGATTAT TTTTAAAGGA 540
 CTCTGTGTAA CACTAACATT TCCAGTAGTC ACATGTGATT GTTTTGTGTT CGTAGAAGAA 600
 TACTGCTTCT ATTTTGAATA AAGAGTTTTT TTTCTTTCTA TGGGGTTGCA GGGATGGTGT 660
 ACACAGAGTC CTAGCATGTA TAGCTGCATA GATTCTTCTA CCTGATCTTT GTGTGGAAGA 720
 25 TCAGAAAGAA TGCAGTTGTG TGTCTATATT TTCCCTCTC AAAATCTTTT AGAATTTTTT 780
 TGGAGGTGTT TGTTTTCTCC AGAATAAAGG TATTACTTTA G

Seq ID NO: 209 Protein sequence

Protein Accession #: XP_059761.1

30 1 11 21 31 41 51
 MALMEVNLIS GFNVPSBAIS LSETVKKVEY DHGKLNLYLD SVNSTQPCVN IPAVRNFKVS 60
 NTQNASVIV DYIEPRRQAV RSYNSEVKLS SCDLCSDVQG CRPCBDGASG SHHSSVIFI 120
 35 FCPKLLYFME LNL

Seq ID NO: 210 DNA sequence

Nucleic Acid Accession #: NM_015472

Coding sequence: 258-1460 (underlined sequences correspond to start and stop codons)

40 1 11 21 31 41 51
 GACACACTCC TCTACAAAC CAGAGACTCC CAAACACAAG GCCTTATATT GACTCATTTC 60
 AGCTCACTATC CTGGCGACTC TCAAGAGAGA AACCTCAGAG TGACTAAATAT CTCCATAATG 120
 45 AAGAAGACATG TACATTCACT ATCTATTTTG GCATTTTCCC CAATACATCT CTGCTCATCT 180
 GACTCTTATC TTGGCATCTG CTTCCTGGTG GATCTGAAC T GACCCATAAG CCAAGCTTAC 240
 TGGTGAATTT CCAGAAGATG AATCCGGGCT CGGCGGCCCC TCCTGCTCCG CGGCTGGGCG 300
 AGCAAGTATC CCAGCTCAGG CAGGACCTAG ACACAGACCT CGAAGCCTCT TTCAACTCTG 360
 50 TCATGAATTC GAAGCTTAGC TCCTGGCGGA AGAAGATCCT GCCGAGTCT TTCTTTAAGG 420
 AGCCTGATTC GGGCTGGCAC TGGCGCCAGT CCAGCACCGA CTGCTGCGGC GGCACCCCGG 480
 GGCTCGACT GGTCTGGGGT GCCCAGCATG TCCGCTCGCA CTGCTGCGGC GGTCTGCTGC 540
 AGCTGGGCAC CGGCGGGGT GCTGCGGTA GCCCGCGCA GCAGCAGCGG CACTCCGCCC 600
 AGCAGTCTCA CGAGTGACC GACGAGCTGC CACTGCCCCG GGGCTGGGAG ATGACCTTCA 660
 CGGCCACTGG CCAGAGGTAC TTCTCAATC ACATAGAAAA AATCACCACA TGGCAAGACC 720
 55 CTAGGAAGGC GATGAATCAG CCTCTGAATC ATATGAACCT CCACCTGCC GTCCAGTTCCA 780
 CACCAATGCC TCAGAGTCC ATGGCAATAT CCCAGCCAAA TCTCTGTATG AATCACCAC 840
 ACCAGCAGCA GATGGCCCCC AGTACCTCTG GCCAGCAGAA CCACCCCACT CAGAACCCAC 900
 CCGCAGGGCT CATGAGTATG CCCAATGGC TGACCACTCA GCAGCAGCAG CAGCAGAAAC 960
 60 TGGCGCTTCA GAGAATCCAG ATGGAGAGAG AAAGGATTGG AATGCGCCAA GAGGAGCTCA 1020
 TGAGGCAGGA AGCTGCCCTC TGTGACAGC TCCCATGGA AGCTGAGACT CTGCCCCCAG 1080
 TTCAGGCTGC TGTCAACCA CCCAGATGA CCCAGACAT GAGATCCATC ACTAATAATA 1140
 65 GCTCAGATCC TTCTCTCAAT GGAGGGCCAT ATCATTOGAG GGAGCAGAGC ACTGACAGTG 1200
 GCCTGGGGTT AGGGTGCTAC AGTGTCCCCA CAATCTCGGA GGACTTCCTC AGCAATGTGG 1260
 ATGAGATGGA TCAGGAGAA AACGCAGGAC AACACCCAT GAACATCAAT CCCCAACAGA 1320
 CCGCTTCTCC TGATTCTCTT GACTGTCTTC CAGGAACAAA CTTTACTTAA GGAATTTTGG 1380
 AATCTGAAGA CCTGATCCCC CTCTTCAATG ATGTAGATGC TGCTCTGAAC AAAAGTGAGC 1440
 CCTTTCTAAC CTGGCTGTAA TCACTACCAT TGTAACCTGG ATGTAGCCAT GACCTTACAT 1500
 70 TTCTTGGGCC TCTTGGAAAA AGTGATGGAG CAGAGCAAGT CTGAGGTGC ACCACTTCCC 1560
 GCTCCATGA TCGTGTCTCC CTCTTTTITA TGTGCCAGT TTAATCATG CCTGGTTTGG 1620
 ATTGAGAGTA ACTTAAGTTA AACATAAATA AATATTCTAT TTTCAITTTT

Seq ID NO: 211 Protein sequence

Protein Accession #: NP_056287.1

75 1 11 21 31 41 51
 | | | | | |

5 MNPASAPPPL PPGQQVIHV TQDLDTOLEA LPNSVMNPKP SSMRKKILPE SPPKPDSDGS 60
 HSRQSSSTDS GHPGPRLAG GAQHVRSRSS PASLQLGTGA GAAGSPAQQH AHLRQQSYDV 120
 TDLPLPPGW EMTFTATGQR .YFLNHIKIKIT TWQDPRKAMN QPLNHNINLHP AVSSTPVPQR 180
 SMAVSQPNLV MNHQQQQMA PSTLSQQNHP TQNPPAGLMS MPNALTQQQ QQQKLRLQRI 240
 QMRERIRMR QELMRQEA LCRQLPMEAE TLAPVQAAVN PPTMTFDMRS ITNNSSDPFL 300
 NGGPHYRSBQ STDGLGLGC YSVPTTPEDF LSNVDEMDTG ENAGQTFMNI NPQQTFFPDF 360
 LDCPLGTNDV LQTLSESDLI PLFNDVRSAL NKSEPLTWL

Seq ID NO: 212 DNA sequence

10 Nucleic Acid Accession #: NM_018174

Coding sequence: 176-2194 (underlined sequences correspond to start and stop codons)

15 CATCTCCCC AACCTGGGG TCGTGTCTT CAACGCTGC GAGCGCGGT CGCGCTGGC 60
 GCGCGCGAG GATGAGCGG AGCTGGCGCT GAGCTCTCT GCGCAGCTG GCATCACGCC 120
 TCTGTCCACT AGCGCGGGCC CCGTGCCAGC CAACCCACC GTGCTCTTG AGAAGATGGG 180
 CTGGGCGCGG CTGGACATGT ATGTGCTGCA CCGCGCTCC GCGCGCGCG AGCGCAGCT 240
 GGGCTCTGTG TGGCGCTGC CCGCGCTGC CCGCGCGCG CCGCGCGAG AGGTGGTGG 300
 CGTCTCTTTC CCGCGTTGCA CCGCGCGCG CTGCTCTCT GAGCGCTGG TCGCTCTGCA 360
 20 GCATCTGAGG TTCTGCGAG AGCCCGTGGT GAGCGCGAG GACTGGAGG GCGCGGGGG 420
 AGCGGAGAGC AAGAGAGAGC TGGGCTCCCG GACAGCTGG AAGAGAGAGG GCCTCTCTGC 480
 CACCCACCTT AGACCTGGCC AGGAGCGCCC TGGGCTGGCC CGCAGGAGC CAGCAGCGG 540
 TGAAGCGCCA CGCAGACTG AGAAGAGAGC CAAGACCCC CGGAGTTGA AGAAGAGCCC 600
 CAAACCGAGT GTCTCCCGA CCCAGCGCG GAGAGTGGC CGGCGAGCT CTCTGTGGC 660
 25 CAACCTCAG AAGAGAGATG CCCAGCGCG ACCCAAGCCC CGCAAGCGC CCAGCAGCTC 720
 CCACTCTGGC TTCCCGCGG TGGCAATGG ACCCGCAGC CCGCGCAGC TCGCATGTGG 780
 AGAAGCGCAG CCGCGCAGT CAGCTCTGG CTCTCGGCC TCCAGCTGG TGGCAGCTC 840
 CAGCTCTGAG CTGGGCGCA TCCAGCGCG GAGGAGAGG GCACTGGAG TGCCTTTGG 900
 CGCGAGCTCA ATCCCAAGC CAGCAGCAG CTCCCTGAG TCCACCGGA GCCCGCGA 960
 30 GCGCGAGAG CCGCTGTGC TGAGCCACT GCGCGCGCG GAGCGCGGC CAGACGCTC 1020
 ACCCAGAGT ACCACACCA CCGTGACAC GCCCTCACTA CCGCAGAGG TGGGCTCCC 1080
 GCACTCAGAC GAGGTGAGC AGTCCCTGTC GTGTCTCTT GAGCAGGTG TGGCGCTATC 1140
 CGCGCCACCC AGTGAAGCTG GCGTGAGCTT CCGCTCTGT CCGCGCGCG GCGCGCGCTC 1200
 GCGCTTCCCA CAGATGTGG ACCTGTGCT GTGTCACTC TGTGAATTT AGCATGCAA 1260
 35 GCGGTGTGCA ATGGCAGC CACCTGCTC CCGCGCAGC TCGAATGAC CAGATGCGC 1320
 GTCACAGGAA CCGCGAGTG GCGTGGGGC CGAGGAGAG CACCCACAT CCGTCAAGCA 1380
 GTCCCTGCCC ACCCTGTCT ACTCGATCC CGTCCCTCT GCGCGCGGT CCGCAGACTC 1440
 AGACGAGAC ACAGAGGGCT TTGGAGTCCC TCGCCAGAC CTTTGGCTG ACCCGCTCAA 1500
 40 GTTCCCGCCA CCACTGCTG ACCCATCCAG CATCTGCAT GTGAGCCCG AGATGCTGCG 1560
 CCGCAGACA GCACGCAAA CGGAGAACGT CAGCGCACC CGGAGCCCG TGGCGCGCC 1620
 CAACTCAGC GCTGCGCGC CCAAGCCAC TCCAGTGGT GCTGCCAAA CCAAGGGCT 1680
 TGTCTGGTGG GACCTGTCCA GCGACCACT CAGTGGCGG AGTGAAGCCA GTGAGAAGG 1740
 AGCGCGGCA CCGCTGTCCA GAAAGTCTC AACCCCAAG ACTGCCACT GAGGCGCTC 1800
 GCGGTGAGC AGCAGCGCG CCGGGGTTC AGCCACCCA CCAAGTCCC CCGTCTACT 1860
 45 GAGCTGCGC TACCTGCCA GCGGAGCAG CGCCACCTG GTGATGAGG AGTTCTTCCA 1920
 GCGCTGTGCG GCGCTGTCT ACCTCATCAG TGGCCAGGAC CAGCGCAAG AGAAGGCAT 1980
 CGCGGCGCTC CTGACGCGC TACTGGCCAG CAGCAGCAT TGGGACCGT ACCTGCAGGT 2040
 GACCGTGATC CCGACTTTG ACTCGGTGG CATGCATAG TGTACGCG AGACGCGC 2100
 50 CCGGCACAG CCGCTGGGA TCACGCTGT GCGCAGCAC GGCATGTGT CCATGCAGGA 2160
 TGACGCTTC CCGGCTGCA AGGTGAGTT CTAGCCCAT CCGCGACAG CCGCGACTC 2220
 AGCCAGCGC GCTGTCTCT AGATTACAG ACATCAGAAA TAAACTGTGA CTACACTTG

Seq ID NO: 213 Protein sequence:

Protein Accession #: NP_060644.1

55 MGVRILMYV LHPSPAGAER TLASVCALLV WHPAGPGEKV VRVLPFGCTP PACLLDGLVR 60
 LQHLRFLREP VVTPQDLEG GRAESKESVG SRDSSKREGL LATHPRPGQE RFGVARKEFA 120
 RAEAPRKTEK RAKTPRELKX DPKPSVSRTO PREVRRAASS VPLNKTNAQ AAPKPKAPS 180
 60 TSHSGPPFVA NGPRSPPSLR CGEASPPSAA CGSPASQLVA TPLSLGPIP AGESEKALELP 240
 LAASSIPRPR TSPSPSHRSP AEGSERLSLS PLRGGAGVD ASPTVTPTV TTPSLPAEVG 300
 SPSTPEVDES LSVSEQVLP PSAPTSEAGL SLPLRGFRAR RSASPHVDLD CLVSPCEPEH 360
 RKAVPMAPAP ASQSSNDSS ARSQERAGL GAEETPPTSV SESLPTLSDS DVPPLAPGAA 420
 DSDSDTEGFG VPRHDPLDF LKVPPLPDP SSICMVDPEM LPPKTRQTE NVSRTRKPLA 480
 65 RPNRAAAPP ATPVAAATK GLAGGDRASR PLGARSEPSE KGRAPLSRK SSTPKTATRG 540
 PSGSASSRPG VSATPPKSPV YLDLAYLPSG SSAILVDEEF PQRVRLCYV ISGQDQKKE 600
 GMRVILDALI ASKQHNDRL QVTLIPTDS VAMHTWYART HARHQAIGIT VLGSNGMVSM 660
 QDDAPFACKV EF

Seq ID NO: 214 DNA sequence

70 Nucleic Acid Accession #: NM_002019.1

Coding sequence: 250-4266 (underlined sequences correspond to start and stop codons)

75 GCGGACACT CTCTCGGCTC CTCCCGGCA GCGCGGCGG CTCGAGCGG GCTCCGGGGC 60
 TCGGTGTCAG CGGCGAGCGG GCCTGGCGGC GAGGATTACC CCGGAGAGTG GTTGTCTCT 120

	GGCTGGAGCC	GCGAGACGGG	CGCTCAGGGC	GCGGGGCGGG	CGGCGGCGAA	CGAGAGGAGG	180
	GACTCTGGGG	GCGGGGTGCT	TGGCGGGGGG	AGCGCGGGCA	CGGGCGGAGC	AGGCGGCGTC	240
	GCGCTCACCA	TGCTCAGCTA	CTGGGACACC	GGGGTCTGTC	TGTGGCGGCT	GCTCAGCTGT	300
	CTGCTTCTCA	CAGGATCTAG	TTGAGGTTCA	AAATTAAGAG	ATCCTGAACCT	GAGTTTAAAA	360
5	GGCACCCAGC	ACATCATGCA	AGCAGGCCAG	ACACTGCATC	TCCAATGCAG	GGGGGAAGCA	420
	GGCCATAAAT	GGTCTTGCC	TGAAATGGTG	AGTAAGGAAA	GCGAAAGGCT	GAGCATAACT	480
	AAATCTGCTT	GTGGAAGAAA	TGGCAACCAA	TTCTGCAGTA	CTTTAACTCT	GAACACAGCT	540
	CAAGCAAAAC	ACACTGGCTT	CTACAGCTGC	AAATATCTAG	CTGTACCTAC	TTCAAGAGAG	600
	AAGGAAACAG	AATCTGCAAT	CTATATATTT	ATTAGTGATA	CAGGTAGACC	TTTGGTAGAG	660
10	ATTGTACAGT	AAATCCCGA	AAATATACAC	ATGACTGAAG	GAAGGGAGCT	CGTCATTCCC	720
	TGCGGGTTTA	CGTCACCTAA	CATCACTGTT	ACTTTAAAAA	AGTTTCCACT	TGCACTTTTG	780
	ATCCCTGATG	GAAACGCGAT	AATCTGGGAC	AGTAGAAAGG	GCTTCATCAT	ATCAATATGA	840
	ACGTACAAAG	AAATAGGGCT	TCTGACCTGT	GAAGCAACAG	TCAATGGGCA	TTTGATATAA	900
	ACAACTATCT	TCACACATCG	ACAAACCAAT	ACAATCATAG	ATGTCCAAAT	AAGCACACCA	960
15	CGCCCACTCA	AATTACTTAG	AGGCCATACT	CTTGCTCTCA	ATTGTACTGC	TACCACTCCC	1020
	TTGAACAGCA	GAGTTCAAAT	GACCTGGAGT	TACCCGTATG	AAAAAATAA	GAGAGCTTCC	1080
	GTAGAGCGAC	GAATTGACCA	AAGCAATTCC	CATGCCAACA	TATCTACAG	TGTTCTTACT	1140
	ATTGACAAAA	TGCAAGAACAA	AGACAAAGGA	CTTTATACTT	TGTGTGTAA	GAGTGGACCA	1200
20	TCATTCAAA	CTGTTAAAC	CTCAGTGCAT	ATATATGATA	AAGCATTCAT	CAGTGTGAAA	1260
	ATCGAAAAAC	AGCAGGTGCT	TGAAACCGTA	GCTGGCAAGC	GGTCTTACCG	GCTCTCTATG	1320
	AAGATGAAGG	CATTTCCTCT	GCGGGAAGTT	GTATGTTTAA	AAGATGGGTT	ACCTGCGACT	1380
	GAGAAATCTG	CTGCTATT	GACTGTTGGC	TACTGTTTAA	TTATCAAGGA	CGTAAGTGAA	1440
	GAGGATGCGA	GGAAATTATC	AATCTTGCTG	AGCATAAJAC	AGTCAAAATG	GTTTAAJAAAC	1500
25	CTCACTGACA	CTCTAATTGT	CAATGTGAAA	CCCCAGATTT	ACGAAAGGCG	CGTGTGATCG	1560
	TTTCCAGACC	CGGCTCTCTA	CCCAGTGGGC	AGCAGACAAA	TCCGTGACTG	TACCGCATAT	1620
	GGTATCCCTC	AACCTACAA	CAAGTGGTTC	TGGCACCCCT	GTAACCAATA	CTATTTCGAA	1680
	GCAAGGTGTG	ACTTTTGTTC	CAATAATGAA	GAGTCTTTTA	TCTGTGATGC	TGACAGCAAC	1740
	ATGGGAACAA	GAATTGAGAG	CATCACTCAG	CGCATGGCAA	TAATAGAAGG	AAAGAATAAG	1800
	ATGGCTAGCA	CCTTGGTGTG	GGCTGACTCT	AGAATTTCTG	GAATCTACAT	TTGCATAGCT	1860
30	TCCATTAAGG	TGGGACTGTT	GGGAAGAAAC	ATAAGCTTTT	ATATCACAGA	TGTGCCAAAT	1920
	GGGTTCATG	TTAACTTGGA	AAAAATGCCG	ACGGAAGGAG	AGGACCTGAA	ACTGTCTTGC	1980
	ACAGTTAACA	AGTTCTTATA	CAGAGACGTT	ACTTGGATTT	TACTGCGAAC	AGTTAATAAC	2040
	AGAACAAATG	ACTACAGTAT	TAGCAAGCAA	AAAAATGCCA	TCACTAAGGA	GCACTCCATC	2100
35	ACTCTTAATC	TTACCATCAT	GAATGTTTCC	CTGCAGATTT	CAGGCACCTA	TGCCCTGCGA	2160
	GCCAGGAATG	TATACACAGG	GGAAGAAATC	CTCCAGAGAA	AGAAATTTAC	AATCAGAGAT	2220
	CAGGAAGCAC	CATACCTCCT	GCGAAACCTC	AGTGATCACA	CAGTGGCCAT	CAGCAGTTCC	2280
	ACCACTTTAG	ACTGTCTATG	TAATGGTGTG	CCGAGGCTCT	AGATCACTTG	GTTTAAAAAC	2340
	AATCAACAAA	TACACACAAG	GCCTGGAATT	ATTTTAGGAC	CAGGAAGCAG	CACGCTGTTT	2400
40	ATTGAAGAGG	TACACAAGAA	GGATGAAGGT	GTCTATCACT	GCAAGAGCCAC	CAACCAAGAG	2460
	GGCTCTGTGG	AAAGTTTCAGC	ATACCTCACT	GTTCAGAGAA	CCTCGGACAA	GCTTAATCTG	2520
	GAGCTGATCA	CTCTAACATG	CACCTGTGTG	GCTGCGACTC	TCTTCTGGCT	CCTATTAAAC	2580
	CTCCTTATCC	GAJAAATGAA	AAGGTCTTCT	TCTGAJATAA	AGACTGACTA	CCTATCAATT	2640
	ATAATGAGCC	CAGATGAAGT	TCCTTTGGAT	GAGCAGTGTG	AGCGGCTCCC	TTATGATGCC	2700
45	AGCAAGTGGG	AGTTTGGCCG	GGAGAGACTT	AJACTGGGCA	AATCACTTGG	AAAGGGGGCT	2760
	TTTGGAAAGG	TGTTCAAGC	ATCAGCATTT	GGCATTAGAA	AATCACTTAC	GTGCGGAGCT	2820
	GTGCTGTGTA	AAATGTCTGAA	AGAGGGGGCC	ACGGCCAGCG	AGTACAAGCG	TCTGATGACT	2880
	GAGCTAATAA	TCTTGACCCA	CAITGGCCAC	CATCTGAACG	TGTTTAACTT	GCTGGGAGCC	2940
	TGCAACCAAGC	AAGGAGGGCC	TCTGATGGTG	ATTGTTCAAT	ACTGCAATAA	TGGAATCTCT	3000
50	TCCAACCTACC	TCAAGAGCAA	ACGTGACTTA	TTTTTTCTCA	ACAAGATGTC	AGCACTACAC	3060
	ATGGAGCCTA	AGAAAGAAAA	AATGGAGCCA	GGCTGGGAAC	AAGGCAAGAA	ACCAAGACTA	3120
	GATAGCGTCA	CCAGCAGGGA	AAGCTTTGCG	AGCTCGGGCT	TTGAGGAGAA	TAAAAGTCTG	3180
	AGTGATGTTG	AGGAAGAGGA	GGATTCTGAC	GGTTTCTACA	AGGAGCCCAT	CATATGGA	3240
	GATCTGATTT	CTTACAGTTT	TCAAGTGGCC	AGAGGCAATG	AGTTCCCTGTC	TTCCAGAAAG	3300
55	TGCAITTCATC	GGGACCTGGC	AGCGAGAAAC	ATTCTTTTAT	CTGAGAACAA	CGTGGTGAAG	3360
	ATTTGTGATT	TTGGCCTTGC	CCGGATATTT	TATAAGAAAC	CCGATTATGT	GAGAAAAGGA	3420
	GATACTCGAG	TTCCCTCTGAA	ATGGATGGCT	CCCGAATCTA	TCTTTGACAA	AATCTACAGC	3480
	ACCAAGAGCG	ACGTGTGGTC	TTACGGAGTA	TTGCTGTGGG	AAATCTTCTC	CTTAGGTGGG	3540
	TCTCCATACC	CAGGAGTACA	AATGGATGAG	GACTTTTGCA	GTGCGCTGAG	GGAAAGCATG	3600
60	AGGATGAGAG	CTCCTGAGTA	CTCTACTCCT	GAAATCTATC	AGATCATGCT	GGAATGCTGG	3660
	CACAGAGACC	CAAAAGAAAG	GCCAGATTTT	GCAGAACTTG	TGGAAAAACT	AGGTGATTTG	3720
	CTTCAAGCAA	ATGTACAACA	GGATGGTAAA	GACTACATCC	CAATCAATGC	CATCTGACAA	3780
	GGAAATAGTG	GGTTTACATA	CTCAACTCCT	GCCTTCTCTG	AGGACTTCTT	CAAGGAAAGT	3840
	ATTTCACTGC	CGAAGTTTAA	TTCAAGGAGC	TCTGATGATG	TCAGATATGT	AAATGCTTTC	3900
65	AAGTTCAATGA	GCTGGAAAG	AATCAAAACC	TTTGAAGAAC	TTTTACCGAA	TGCCACCTCG	3960
	ATGTTTGATG	ACTACCAAGG	CGACAGCAGC	ACTCTGTGCG	CCTCTCCCAT	GCTGAAGGUC	4020
	TTCACTCTGA	CTGACAGCAA	ACCCAGGGCC	TGCTCAAGAA	TTGACTTGA	AGTAACCACT	4080
	AAAAGTAAGG	AGTCGGGGCT	GTCTGATGTC	AGCAGGCCCA	GTTCCTGCCA	TTCCAGCTGT	4140
	GGGCACTGCA	GCGAAGGCAA	GCGCAGGTTT	ACCTACGACC	ACGCTGAGCT	GGAAAGGAAA	4200
70	ATGCGCTGCT	GCTCCCGGCC	CCCAGACTAC	AACCTGGTGG	TCTGTACTCT	CACCCACCC	4260
	ATCTAGAGTT	TGACACGAAG	CCTTATTTCT	AGAAGCACAT	GTGTATTTAT	ACCCCAAGGA	4320
	AACTAGCTTT	TGCCAGTATT	ATGCATATAT	AAGTTTACAC	CTTTATCTTT	CCATGGGAGC	4380
	CAGCTGCTTT	TTGTGATTTT	TTTAATAGTG	CTTTTTTTTT	TTGACTAACA	AGAAATGTAAC	4440
	TCCAGATAGA	GAATAGTGA	CAAGTGAAGA	ACACTACTGC	TAAATCCTCA	TGTTACTCAG	4500
75	TGTTAGAGAA	ATCTTCTCTA	AACCCAAATGA	CTTCCCTGCT	CCAAACCCCG	CCACTCTCAG	4560
	GACGCAAGGA	CCAGTTTGAT	TGAGGAGCTG	CAGTGATCAC	CCAATGCATC	ACGTACCCCA	4620
	CTGGGCCAGC	CCTGCAGCCC	AAAACCCAGG	GCAACAAGCC	CGTTAGCCCC	AGGGGATCAC	4680

Seq ID NO: 215 Protein sequence:
Protein Accession #: NP_002010.1

55	1	11	21	31	41	51	
	MVSYNDTGVL	LCALLSCILL	TGSSSGSKLK	DPELSJKCTQ	HIMQAGQTIL	LQCRGEAAHK	60
	WSLPEMVSKR	SERLSITKSA	CGNGKGQFCS	TLTLNTAQAN	HGTFGSKCYL	AVPTSKKQBT	120
60	ESAIYITFSD	TGRPFVMTYS	BEPILHMTS	GRSLVPCRV	TSNPITVTLK	KPFLDTLIFD	180
	GKRIIWDSCR	KPIGNAATYK	BIGELTCBAT	VNGHLVYKTL	LTHRQNTWII	DVQISTPRPV	240
	KILLRGHTVL	NCATTSPFAT	RVQMTVSYSP	BKKNRKSVR	RIDQSLVHAN	IFYSVLTIDK	300
	MONKDKGLT	CRVSGSPSPK	SVMTSVHIVD	KAPITVKHR	QQVSLTAVGK	RSYRLSKMVK	360
	APPSBPVWVL	KDGLPATERK	ARYLIRGYSL	ILKDVTEBDA	GNVYTLITKL	QSNVFNKTLA	420
65	TLVNVKPSI	YKAVSSFPD	PALYPLSPKQ	ILCTCAVPI	OPTTKMFWP	CHNHSEARC	480
	DFCSNNESP	LIDADSNMGN	RIESITQRKA	IIKGKSTVNI	TLVVADSRIS	GIYICIAENK	540
	VTGVRNHSI	YITDVPFHG	VNLKMKPTGS	BDLKLCTSVN	KFLYIRDVTI	LKRTVNRWTK	600
	HYISKQKQMA	ITKHSITGL	LTIMVSLTQ	SGTYACRAN	VYTGRIILQK	KEITIRDEQA	660
70	PYLRNLSVDH	VTASISSTLT	DCHANGVPEP	QITWFQNNH	IQQRQPIFL	PGSSTLFTFR	720
	FTLDEBGVTH	CKATNGKSGV	ESSAYLTVQK	TSDKSLMELK	LTITCTVAAI	PLFWLLTLTI	780
	RKMKRSSPBI	KDYLISLIKD	PDEVLPEQAC	ERLPLVDAEK	EPARERLIKG	KSLRGAGPCK	840
	VVQASAPGKI	KSPCTRTVAV	KMLKREGATS	EYKALMTBKL	LITHLGHVN	VNLLGACAT	900
	QGGPLMVIHV	YCKYKNSLSY	LKSKRDLFFL	NKDAALMEDP	KKSKMPEGLS	QQKPKRLDSV	960
75	TSRSSFASGG	PQCKGSLSDV	EBESDSGFL	KEPITLMEPI	SPYSQVAPR	EFLSRKSRST	1020
	RDLAARNLNI	SSNNVVKICD	FGLARDIYKN	PDYVRIGMTR	LEPLKNWAPES	IFDKIYSTKS	1080
	DVMSYGVLLM	BIFLSSGQSL	PVGOMDEDFP	SRLESGMRDR	AEYSTPYTQI	QIMLDCKWRT	1140

PKERPRFAEL VEKLGDLQA NVQODGKDYI PINAILTGNS GFTYSTPAPS EDFFKESISA 1200
 PKFNSGSSDD VRYVNAFKPM SLERIKTFEE LLFNATSMFD DVQDSSTLL ASPMLKRFTW 1260
 TDSKPKASLK IDLRVTSKSK ESGLSDVSRP SPCHSSCGHV SBGKRRTYD HAGLERKIAC 1320
 CSPPPDYNSV VLYSTPPI

5

Seq ID NO: 216 DNA sequence

Nucleic Acid Accession #: NM_024689

Coding sequence: 76-624 (underlined sequences correspond to start and stop codons)

10 1 11 21 31 41 51
 CTCTTTGGCC AAGCCCTGCC TCTGTACAGC CTGAGTGGG CAGCCAGAGG CTGCAGCTGG 60
 AGCCAGAGCC CCAAGATGGA GCCCCAGCTG GGGCCGTGAG CTGCCGCCCT CGCCCTGGC 120
 TGGCTGGCCC TGCTGCTGGT GGTCTCAGCC CTGAGCTGTT CTTCTCTCTT GCCAGCTTCT 180
 15 TCCCTTTCTT CTCTGGTGGC CCAAGTCAGA ACCAGCTACA ATTTTGGAG GACTTTCTCT 240
 GGTCTTGATA AATGCAATGC CTGCATCGGG ACATCTATT TGAAGAGT CTTTAAAGAA 300
 GAAATAAGAT CTGACAACTG GCTGGCTTCC CACCTTGGAC TGCCTCCCGA TTCTTGCTT 360
 TCTTATCCTG CAAATTAATC AGATGATTCG AAAATCTGGC GCCCTGTGGA GATCTTTAGA 420
 CTGGTCAGCA AATATCAAAA CGAGATCTCA GACAGGAAAA TCTGTGCTTC TGCATCAGCC 480
 20 CCAAGAGCCT GCAGCATGGA GCGTGTCTGT CGGAAAACAG AGAGGTTCCT GAAATGGCTG 540
 CAGGCCAAGC GCCTCAGGCC GAGCTGTGTT CAGGACTGTG ACCAGGGCCA GAGAGAACTA 600
 AAGTCTCTGT GTATGCTGAG ATAAACCCAG TGAAGAAAGC TGGCATGGAG CCGACCATG 660
 AGAACTTCCA GAAAGTGTTA GCCTTCTCCC AACTGTGTTA TACCAACCAT ATTTTCAAA 720
 AGTAATCAAT AAGAGGCTT CTGCATCAAA CCTTCACATG CAGCTCCCAT GCCACCTCC 780
 25 AGAATTCACC AACACACAGG CCCACACGCA ACAGGCTACC TTTCACAAAT ATTCCTGTAT 840
 GACAACTCCA AAGCCCGGCC TCTTTCCACC ACACGTGTGT CCGCTAGATG GGGCTGTGTC 900
 TGAGCCCAAC CCAATCCAGA TGTGATCCCC CTGTGATCTA CTTCTGGCAA GATTCTCAGT 960
 CTGACAGGTT TCCCTATG AGATAGAACC TGATAAGGAG CTAGGCAAT TCTGACAACT 1020
 TTACCAAGG CCAACATAAC TTCTAAATTT TGGTCTGGTC TGAAGGAAAA CCGTCTCTCG 1080
 30 CCCTAGTGTG GAGTGAACCT TCTTATCTCT GGCTTCTAGA GGGAAAAAAA AAGCATACCT 1140
 CTTTACTTTT TAAAGTAACT CCATCAGAGT CATGAATACA CCGTCAAGA CTATCTATCT 1200
 TTTATGTTTC CATCTGGTGA AGAATCTCTT AAATGAGGAC ACTGCTGATT GCTGGTGATG 1260
 TTTTGTGAGC AAACACTGGG GGGTATGGAT GAAAGCCCAT CCGAGGTCAA ATGACTCCTT 1320
 GGGGAAGCTA CTCTCTCTCT ATTCAATTT CACTAAATC TTCCAAGATG AAAGCAATC 1380
 35 TAGATTTCCG TCTTCAATGC TGTCCATTTT TGTAAATGAC GAGTGTTTTT CTTTAGCTA 1440
 GTGTATCAGG CAGGTTCTTA CCAGAGAAAC AGAACCACTA GGAGATACAT ATACATGTCC 1500
 AGATTTATTT CAAAGAAATG ATTTACATGA TTGTGGGAT TGGCAAGTCG AAAATCCATA 1560
 TGGTAGGCTT GCAATCTGTA AACCTTTGGG CAGGAGCTGA TGGTGTAGTT TGCAGATAGA 1620
 40 ATCTCTTGTT CTTTAAAAAA ATCTGTCTTT GTCTTAAAG GCTTTGAATG ATTTGATCAG 1680
 GCCCACCAGC ATTACCTAGA TAATCTCTTT TACTTAAAGT AAATGATG TAGGTGCTAA 1740
 TCACATCTAT GAAATGCCCT CACAGCAACA CCTAGATTAG CATTCAATTG AATAACTGGG 1800
 GAATACAGCC TAGCAGATT GACACATAAA ATTAACCATC ACAGCAACAT GCCTGTCTAA 1860
 TTTTATGTTT CATCTTCAGA CTGTTAAGGA TTGTGTGAGA GAACTGTGAC AGCACTCTC 1920
 45 AGCATCACCC TGAACCAAG GCCCTATCA AGTAACAATA TAGCCAAGCA AAATTCAGT 1980
 CAATAGAGAC ATTGACTGGT TGGCTGGCTT CCAAGGGAT AGCAACGAC AAGAAATGCA 2040
 AGGATAGGGA AACCAAGCAC GGGAGAGGGA GGGGCAACAG AGGTCCAGGG TTTGTTTATC 2100
 TTTTATTTT TCACTGGGAG GTGGTAAATT AGCCCTGTTT CCAATGTATG CAGATGGGAG 2160
 AAGTGAATTA GAAATCCAA AGCAATTTGT AATCCCAAAA ATGGGTGTAT CTGGTTTGA 2220
 50 ATGAACCTT ATTTATATG AAATGGTTGG TTCCCAATT CTGTTTGCCA TTGGCCAATA 2280
 TAATTTGGG TTTGCAATG GCCAGCACAT GCCAAACAGA AGTAGACAAA GGTCTCACTC 2340
 TGTAAAGTGG ACCTTGGGGA GGAGCTGGCT CCATCATAAA GGGAGGGGTT AGTAAAAATG 2400
 GTCTCTTAAG CCTGTCTCTG CTACAGTTAT AGAGGTTGCT CAGAACCTTC TCAGCAATA 2460
 TAGCAGTTAT CTATTGTTGT GTATTAAACC ATTTCAACAC AT

55

Seq ID NO: 217 Protein sequence

Protein Accession #: NP_078965.1

60 1 11 21 31 41 51
 MEPLQGPBAA ALRPGMLALL LWSALSCSF SLPASSLSLL VPQVRSYNF GRTPLGLDKC 60
 NACIGTSICK KPFKEBIRSD NWLASHLGLP PDSLLSYFAN YSDSKIWRF VEIPLVSKY 120
 QNEISDRKIC ASASAPKTC IERVLKTER PQRLQAKRL TPDLVQDCHQ GQRELKFLCM 180
 LR

65

Seq ID NO: 218 DNA sequence

Nucleic Acid Accession #: AF075027.1

Coding sequence: 3-269 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 GATTAAATTAA GTGCTTTAAA CCGTCTTGGT AAATATTCG CCGGAGCTGG GGAGGACCGT 60
 TGGGATGGCT GTAGCTTGAG TTGAATTTTA ACTGTCTCTA TTCTGGGTTT TGTGCTCTG 120
 CTTTCTGTGC CAGGTGCTG TGTACGGGA GAGAGTGACT GGAAGTAAC AAAGCTGAAT 180
 75 CTTTCTCCCT GGAGTAAGGC CGAAGACTGG ATTACTACAC GCCTAGACGT GACACTACAC 240
 CCATAGATCT CATGCATCAT TAATGCCATA TGACATTGCC ATTTTCTTTC TCAGTTCAG 300
 GACAAAAGTG GTGGGTTTTC ATTTCTTCA CTGATTGTCA ATGCATTAA TAAAGAAAGAT 360

TGTGGT

Seq ID NO: 219 Protein sequence
Protein Accession #: AF075027

1	11	21	31	41	51	
ERKMQCHMAL	MMHEIYGC	TSRRVVIQSS	ALLQGERPSP	VTQSLSPVT	QHLGTSRAT	60
KPRMRTVKIQ	LKLQPSQRSS	PAPAEYLPRP	FKALN			

Seq ID NO: 220 DNA sequence

Nucleic Acid Accession #: AL133411.8

Coding sequence: 1-1395 (underlined sequences correspond to start and stop codons)

15	1	11	21	31	41	51	
	<u>ATGGGCAAGG</u>	ACTTCATGAC	TAAAACACTA	AAAGCAATGG	CAACAAAAGC	CAAAATTGAC	60
	AAATGGGATC	TAATCAAATT	AAAGAGCTTC	CGCACAGCAA	AAGAACTAT	TATCAGAGTG	120
	AACAGGCAC	CTACAGAATG	GGAGAAAAT	TTTGCAATGT	ATCCATCTGA	CAAGGGCTG	180
20	ACATCCAGAA	TCTATAAGGA	ACTTAAACAA	TTTTACAAGA	AAAAACCAA	CAAGCCATC	240
	AAAAAGACA	TGGATGAAGC	TGGAAACCGT	CATTCTCAGA	AAACTAACAC	AGGAACGAA	300
	AACCAACAC	CACATGTTCT	CACATCATAG	TGGGAGTTGA	ACAATGAGAA	CACATGGGA	360
	CAGGGAGGG	AACATCACAC	ACTGGGGCTT	GTGAGAGGCC	CCTCTGGCT	CCTGGCTGGC	420
	CTTGAACATG	CTGGGAGGAA	ATTACAAATC	ATCCATGGGC	TGTTTACCT	TGAAAATGAA	480
25	TGGGCCAGG	AACATCCAT	AATACAAAG	AAATATGCAT	TATGGATTGG	AAACCAAGCAG	540
	ATCTGGGTGG	CACAACTCC	TGGTGAATCT	ATCTCCAGTT	CACCAGCATT	GCCTAATGTG	600
	CTACCTTTAA	ATGAAGATGT	TAATAAGCAG	GAAGAAAGA	ATGAAGATCA	TACTCCCAAT	660
	TATGCTCCTG	CTAATGAGAA	AAATGGCAAT	TATTATAAAG	ATATAAACA	ATATGTTGTC	720
	ACAACACAAA	ATCCAAATGG	CAGTGAATCT	GAATATCTG	TGAGAGCCAC	AACAGACCTG	780
30	AATTTTGTCT	TAAAAAAGA	TAAACTGTG	AATGCAACTA	CATATGAAAA	ATCCACCAT	840
	GAAGAAGAAA	CAACTACTAG	CGAACCTCTT	CATAAAATA	TTCAAGATC	AACCCCAAC	900
	GTGCTGTCAT	TTTGACAAT	GTAGCTAAA	GCTATAAATG	GAACAGCAGT	GGTCAATGAT	960
	GATAAAGATC	AATTATTCA	CCCAATTCCA	GAGTCTGATG	TGAATGCTAC	ACAGGGAGAA	1020
	AATCAAGCAG	ATCTAGAGGA	TCTGAAGATC	AAAAATATGC	TGGGAATCTC	GTGATGACC	1080
35	CTCCTCTCTT	TTGTGTTCTT	CTTGGCATTG	TGTAGTGCTA	CAGTGTACAA	ACTGAGGCAT	1140
	CTGAGTTATA	AAAGTTTGA	GAGTCAATAC	TCTGTCAACC	CAGAGCTGCG	CACGATGCTT	1200
	TACTCTCATC	CATCAGAAGG	TGTTTCAGAT	ACATCTCTTT	CCAAGAGTGC	AGAGAGCAGC	1260
	ACATTTTGGG	GTACCACTTC	TTCCAGATATG	AGAAGATCAG	GCACAGAAGC	ATCAGAACTT	1320
40	AGATAATGA	CGGATATCAT	TTCCATAGGC	TCAGATAATG	AGATGCATGA	AAACGATGAG	1380
	TCGGTTACCC	<u>GGTGA</u>					

Seq ID NO: 221 Protein sequence
Protein Accession #: AL133411.8

45	1	11	21	31	41	51	
	MCKDFMTKYL	KAMATKAKID	KNDLIKLSF	RTAKETIIRV	NRQPTWEKYN	FAMYPSEKGL	60
	TSRIYKELKQ	PKKKPNNAI	KKMDRAGNR	HSQKTMGTGE	NQTPHVLTHK	WELANMENTWT	120
50	QCGSHHTLQP	VRSPSGLLAG	LEHAGRKLP	IHLFTLENE	WAGQSIYQK	KYALWIGTKQ	180
	IWVAQTPOBS	IGSSPALPNV	LPLNSDVNKG	BEKNEDHTFN	YAPANEKNGN	YKDKIQYVVF	240
	TIQNPNGTES	SISVRATDGL	NFALKNDKIV	NATTYKASTI	BSSTTTSBPS	HKNIQRSTFN	300
	VPFWMPLAK	AINGIAVMD	DKDQLPHPI	ESDVNATQGE	NQPDLEDLKI	KIMLGISLMT	360
55	LLLFVLLAP	CSATLYKLKH	LSYKSCSESY	SVNPELATMS	YFHPSEGVSD	TSFSEKSESS	420
	TFLOTTSSDM	RRSGTRTSBS	KIMTDIISIG	SDNEMHENDE	SVTR		

Seq ID NO: 222 DNA sequence

Nucleic Acid Accession #: AL050295.1

Coding sequence: 237-2073 (underlined sequences correspond to start and stop codons)

65	1	11	21	31	41	51	
	GAAGGGGACA	GAAGGCAGTT	CACCTCTGCT	CCCGACAGCC	TGGGAACCGG	CAAGAGCCCC	60
	AGCATTTGAA	GTCTGGTCTT	GTGAAACCCC	ACCCTCCTCT	GGCTGTGTGA	TTGAATGGGA	120
	TGCCCTGAGC	GTACACCTCA	CTGAGAGGGG	TTTGGGSCAG	ATCAGCAGTA	AGGTGTTAAA	180
	TTTTTGAAGC	CTGAAACTC	CAGAAGAGAA	AGGCCAACCA	ACTCAAACCT	GAAGACATGA	240
	AATCCCAAG	GAGAACCACT	TGTGCTCTCA	TGTTTATTGT	GATTTATTCT	TCCAAGCTG	300
70	CAGTGAAGTG	GAATTACGAG	TCTACTATTG	ATCCTTTGAG	TCTTCATGAA	CATGAACCA	360
	CTGTTGAAGA	GGCACTGAGG	CAAAAACGAG	CGTTTGGCC	AAAAAGTCTT	ACGGCTGAAG	420
	AATACACTGT	TAATATTGAG	ATCAGTTTGG	AAAAATGCAT	CTTCTGGAT	CCTATCAAG	480
	CCTACTTTGAA	CAGCCTCAGT	TTTCCAATTC	ATGGGAATAA	CAGTGAACCA	ATTACTGACA	540
	TTTTTGAAGC	AAATGTGACA	ACAGTCTGCA	GACCTGCTGG	AAATGAAATC	TGGTGTCTCT	600
75	GCGAGACAGG	TTATGGGTGG	CCTGGGGAAG	GGTGTCTTCA	CAATCTCATT	TGTCAAGAGC	660
	GTGAGCTCTT	CCTCCAGAGG	CACCATTTGA	GTTCCTTTAA	AGAACTGCTT	CCCAATGGAC	720

CTTTTCGCTT GCTTCAGGAA GATGTTACCC TGAACATGAG AGTCAGACTA AATGTAGGCT 780
 TTCAAGAAAG CCTCATGAAC ACTTCCTCOG CCCTCTATAG GTCTACAG AGCGACTTGG 840
 AAACAGCGTT COGGAAGGTT TACGGAATTT TACCAGGCTT CAAGGGCGTG ACTGTGACAG 900
 GGTTCAGATC TGAAGTGTG GTTGTGACAT ATGAAGTCAA GACTACACCA CCATCACTTG 960
 5 AGTTAATACA TAAAGCCAAT GAACAAGTTG TACAGAGCCT CAATCAGACC TACAAAATGG 1020
 ACTACAACCTC CTTTCAAGCA GTTACTATCA ATGAAAGCAA TTTCTTTGTC ACACAGAAAA 1080
 TCATCTTTGA AGGGGACACA GTCAGTCTGG TGTGTGAAA GGAAGTTTTC TCCTCCAATG 1140
 TGTCTTGGGG CTATGAAGAA CAGCAGTTGG AAATCCAGAA CAGCAGCAGA TTCTCGATTT 1200
 ACACGCCACT TTTCAACAAC ATGACTTCGG TGTCCAAGCT CACCATCCAC AACATCACTC 1260
 10 CAGGTGATGC AGGTGAATAT GTTGTCAAA TGTATATTAGA CATTTTTGAA TATGAGTGCA 1320
 AGAAGAAATAT AGATGTTATG CCCATCCAAA TTTTGGCAA TGAAGAAATG AAGGTGATGT 1380
 CGCAATGCAA TCCTGTATCT TTGAAGTCTT GCAGTCAGGG TAATGTTAAT TGGAGCAAGG 1440
 TAGAATGGAA GCAGGAAGGA AAAATAAATA TTCCAGGAAC CCCTGAGACA GACATAGATT 1500
 CTAGCTGCAG CAGATACACC CTCAGGGCTG ATGGAACCCA GTGCCCAAGC GGGTCTGCTG 1560
 15 GAACAACAGT CATCTACACT TGTGAGTTCA TCAGTGCTTA TGGAGCCAGA GGCAGTGCAA 1620
 ACATAAAAGT GACATTTCATC TCTGTGGCCA ATCTAACAA TAAACCCGAC CCAATTTCTG 1680
 TTTCTGAGGG ACAAAACCTT TCTATAAAAT GCATCAGTGA TGTGAGTAAC TATGATGAGG 1740
 TTTATTGGAA CACTTCTGCT GGAATTAAAA TATACCAAG ATTTTATACC ACGAGGAGGT 1800
 ATCTGTATGG AGCAGAATCA GTACTGACAG TCAAGACCTC GACCAGGGAG TGGAAATGGAA 1860
 20 CCTATCCTAG CTATTTTGA TATAAGAATT CATACAGTAT TGCACCAAAA GACGTCATTG 1920
 TTTCAACGCT CCCTCTAAG CTGAACATCA TGTATTGATCC TTTGGAAGCT ACTGTTTCAT 1980
 GCGATGTTTC CCATCACAAC AAGTGCTGCA TAGAGGAGGA TGGAGACTAC AAGTTACTT 2040
 TCCATATGGG TTCTCATCC CTTCCTGCTG TAAAAA AAAAAAAAAA A

Seq ID NO: 223 Protein sequence
 Protein Accession #: CAB43394.1

1 11 21 31 41 51
 30 MKSPRRRLTL LMFIVIVSSK AALMWNVYEST IHPLSLHSEH PAGEALRQK RAVATKSPTA 60
 EETVNIIEIS FENASFLDPI KAYLNSLSFP IHGNVTDQIT DILSINVTV CRPAGNEINC 120
 SCRTYGVNPR ERCLHNLICQ ERDVLPGHH CSCLKLPPN GPFCLLQEDV TLMNRVRLNV 180
 GFGDLMMTS SALLYRSYKTD LETAFLKGYG ILPGFKGVTV TGFKSQSVV TVEYKTTTTPS 240
 35 LELIHKANEQ VVQSLNQTYK MDYNSPQAVT INESNPFVTP BIIPEGDTVS LVCEKEVLS 300
 NVSWRYBEGQ LELQNSRRFS IYDALPNNMT SVSKLTIHNI TPGDAGEYVC KLILDIFEYE 360
 CKKKLDVNP IILANBEMKV MCDNNPVSIN CCSQGNVNS KVENKQBGKI NIPGTPETDI 420
 DSSCSRITLK ADGTQCPGGS SGTTVIYTCR FISAYGARG ANIKVTFISV AMLITTPDI 480
 SVSEGNFPI KCTSDVSNYD EYVWNTSAGI KIYQRYFTR RYLDGAEVSL TVKTSTREWN 540
 40 GTVHCIFRYK NYSYIATKDV IVHPLPLKLN INIDPLEATV SCSSGSHIKC CIEEDGDYKV 600
 TFMGSSSLP AVKKKKKK

Seq ID NO: 224 DNA sequence

Nucleic Acid Accession #: NM_007268

Coding sequence: 46-1245 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 45 GGTAGCAGGA GGTGGAAGA AAGGACAGAA GTAGCTCTGG CTGTGATGGG GATCTTACTG 60
 GGCCTGTACT TCTTGGGGCA CTAACACAGT GACACTTATG GCCTGCCCAT CCGTGAAGTG 120
 50 CCAGAGAGTG TAACAGGACC TTGGAAGGGG GATGTGAATC TTCCCTGCAC CTATGACCCC 180
 CTGCAAGGCT ACACCCAAAG CTGTGTGAAG TGCGTGGTAC AACGTGGCTC AGACCTCTGC 240
 ACCATCTTTC TACGTGACTC TTCTGGAGAC CATATCCAGC AGGCAGAGTA CCAGGGCCGC 300
 CTGCATGTGA GCCACAAGGT TCCAGGAGAT GTATCCCTCC AATTGAGCAC CCGTGGAGTG 360
 GATGACCGGA GCACTACAC GTGTGAAGTC ACCTGGCAGA CTCTGTATGG CAACCAAGTC 420
 55 GTGAGAGATA AGATTACTGA GCTCCTGTGC CAGAACTCT CTGTCTCCAA GCCCACAGTG 480
 ACAACTGGCA CGGTTATGG CTTCAAGGTG CCCCAGGGAA TGAGGATTAG CTTCAATGTC 540
 CAGGCTCGGG GTTCTCTCC CATCAGTTAT ATTTGGTATA AGCAACAGAC TAATAACAG 600
 GAACCCATCA AAGTAGCAAC CTAAGTACC TTACTCTTCA AGCCTGCGGT GATAGCCGAC 660
 TCAGGCTCCT ATTTCTGCAC TGCCAAAGGC CAGGTGGCTC CTGAGCAGCA CAGGACATT 720
 60 GTGAAGTTTG TGGTCAAAGA CTCTCAAAG CTAAGTCAAG CCAAGACTGA GGCACCTACA 780
 ACCATGACAT ACCCCTTGAA AGCAACATCT ACAGTGAAGC AGTCTGGGA CTGAGCCACT 840
 CAGATGAGTG GCTACCTTGG AGAGACCAAT GCTGGGCCAG GAAAGAGCCT GCCTGTCTTT 900
 GCCATCATCC TCATCATCTC CTGTGTCTGT ATGGTGGTTT TTACCATGGC CTATATCATG 960
 65 CTCTGTCCGA AGACATCCCA ACAAGAGCAT GTCTACGAAG CAGCCAGGCG ACATGCCAGA 1020
 GAGGCCAACG ACTCTGAGA AACCATGAGG GTGGCCATCT TCGCAAGTGG CTGCTCCAGT 1080
 GATGAGCCAA CTTCCAGAA TCTGGGCAAC AACTACTCTG ATGAGCCCTG CATAGGACAG 1140
 GAGTACCGA TCATCGCCCA GATCAATGGC AACTAGGCCG GCCTGCTGGA CACAGTTCTT 1200
 CTGGATTATG AGTTCTTGGC CACTGAGGCG AAAAGTGTCT GTTAAAAATG CCCCATTAGG 1260
 70 CCAGGATCTG CTGACATAAT TGCCATAGTCA GTCCCTTGCT TCTGCATGAC CTCTTCTCCT 1320
 GCTACCTCTC TTCTCGGATA GCCCAAAGTG TCCGCTTACC AACCTGGAG CCGCTGGGAG 1380
 TCACTGGCTT TGCCCTGGAA TTTGCCAGAT GCATCTCAAG TAAGCCAGCT GCTGGATTG 1440
 GCTCTGGGCC CTTCTAGTAT CTCTGCCGGG GGCTTCTGCT ACTCTCTCT AAATACCAGA 1500
 GGAAGATGTC CCATAGCACT AGGACTTGGT CATCATGCTC ACAGACACTA TTCAACTTTG 1560
 75 GCATCTTGCC ACCAGAAGAC CCGAGGGAGG CTCAGCTCTG CCAGCTCAGA GGACAGCTA 1620
 TATCCAGGAT CATTTCTCTT TCTTCAGGGC CAGACAGCTT TTAATTGAAA TTGTTATTTC 1680
 ACAGGCCAGG GTTCAGTTCT GCTCCTCCAC TATAAGTCTA ATGTTCTGAC TCTCTCTG 1740

WO 02/079492

PCT/US02/04915

TGCTCAATAA ATATCTAATC ATAACAGCAA AAAAAAAAAA AAAAAA

Seq ID NO: 225 Protein sequence:
Protein Accession #: NP_009199.1

5
10
15

1	11	21	31	41	51	
MGILLGLLLL	GHLVDVYGR	PILEVPFSVT	GPWKGDVNL	CTYDPLQGYT	QVLVXWLVR	60
GSDPVTIFLR	DSSGDHIQQA	KYQRLHVS	KVPGDVSLLQ	STLMDDRSH	YTCEVTWQTP	120
DGNQVVRDKI	TLRLVQKLSV	SKPTVTGSG	YGFTVPQGR	ISLQCAQSG	PPISYIWKQ	180
QTNMQEPIKV	ATLSTLLFKP	AVIADSGSYF	CTAKGVGSG	QHSDIVKPVV	KDSSKLLKTK	240
TRAPTMTYTP	LKATSTVKQS	WDWTTDMG	LGSTAGVPGK	SLPVFAILLI	ISLCCMVVPT	300
MAYIMLCRKT	SQSEHYVEAA	RAHAREANDS	GETNRVAIPA	SGCSSDEPTS	QNLGNNYSDE	360
PCIGQBYQII	AQINGNYARL	LDTVPLDYEF	LATEGKSV			

Seq ID NO: 226 DNA sequence

Nucleic Acid Accession #: XM_64321

Coding sequence: 1-2079 (underlined sequences correspond to start and stop codons)

20
25
30
35
40
45
50
55
60

1	11	21	31	41	51	
ATGTCGCCA	GTTCCGATCA	AGACAGAGCC	CGGTATCTTC	CAGGACACT	AGACAAGATG	60
CCAGAACAC	GCTCCGCTC	TGCCAGAGG	CCAAAGCAG	CCCAACAAGA	GCCCGGCATT	120
GAGCCCTGTA	CTTACAGGA	GGTGTGTGA	GCCATCGTCC	TCACTGATGC	GCTCGGGATC	180
GGGTTTGGGA	TCACGGGAAA	CACAGTTCAA	CAACCACTC	AACTCACTGA	CTCGCCAGC	240
ATCCGTCAAG	AGGATGCTT	TGATAACAAA	ATTGACATTG	CTGAAGATGG	TGGCCAGACA	300
CCATACAAAG	CTACCTTGCA	GCAAGCTTT	CAATACTCAC	CTACAAAGA	TCTTCTCCA	360
CTACAAATG	GCTACCTGCC	ATCAATCAGC	ATGTATGAAA	TTCAAACCAA	ATACCAGTGG	420
CTATATCAAT	ATCTATATGG	AAATTCATAA	CAGAAGACCA	CATTAAATTC	TAGAAAACCC	480
TTCCCTCCCA	CAGCCACCAC	TTGGTATCCA	CAAACTGTGA	TTCAAAGAA	GAGTGGCTCA	540
CCTGAAGTTA	AACTAAAAAT	AACCAAACT	ATCCAGAATG	GCAGGGAAAT	GTTCAAGTCT	600
TCCTTTTGTG	GAGACCTTTT	AAATGAAAGT	CAGGCAAGTG	AGCACACGAA	GTCAAAGCAT	660
GAAAGCAGAA	AAGAAAGAG	GAAAAAACCC	AAAAAGCATG	ACTCATCAAG	ATCTGAAGAG	720
CGCAAGTCAC	ACAAAATCCC	CAAAATTAGAA	CCAGAGGAAC	AAAAATAGACC	AAATGAGAGG	780
GTTTACACCA	TATCAGAAAA	ACCAAGGGAA	GATCCAGTAC	TAAAGAGGGA	AGCCCCAGTT	840
CCCAACATAC	TATCTTCTGT	TCCAACAACA	GAAAGTGTCA	CTGTGTATA	GTTTCAAGTT	900
GGTGATCTTG	TGTGTCCAA	GGTGACGGTC	ACACCTGTGT	GGTGTCCCGC	CTGTGAGGGA	960
CGGAGGAGCC	ATCACTGTTC	CAGCTGCTTG	GAGATCTTGG	TGCTGTGTGC	AGCCCTCAGC	1020
CTCAAGAGGT	CTTCTCTGTT	TTCTTCTCTG	AGTTCCTCTA	CCTCCACGGG	CAACACAGAG	1080
CTCAAGATCT	AGGGAACCTG	CCAGATGGGC	TGCTCACTTA	TGGCTCTCAC	GACCAATGTC	1140
TCCTCTCTCC	TGTGTCATTG	GGAAGGAACA	GACCAGATGT	CATCCAGGGG	CCCGGAATTT	1200
GGGGGGGCGC	GCTGGGTGTG	GCAGCATCAG	AAGCCTCAGA	TCCGCACTCT	CATCTGCCAC	1260
AGGCCAGGGA	AGGAACCTCT	GAGACTCAGT	TTCTTACGAT	GTAAGATGGA	GAGAAGAATC	1320
TCCTCTTTAG	CCACCTCTCA	GGCTGCTGCG	TGTTCCGCCC	CAGACCACTG	CTTGTGAGAA	1380
TGCTTATAGAG	ACTATGCAAG	GGGCGGCCAT	TTGACACTCA	GAGCCACGGA	AGCCTTTCTT	1440
GGTCCAGACA	GCAGGACTCG	AGGCTTCTAG	GCTGTGCGCA	AGAGATACTG	CAGGAACAGC	1500
CAGCACCAGA	GATATCTCCT	GCAAGGCCTC	CTAGGTGGGT	TCTTGAAGA	AAGGAATGCC	1560
AATGAATATG	ATTGCAAGCT	AGAGACGAGA	GAAGCGCGGT	CCTCAACTCC	AAGAAATCCG	1620
TATTCCCCAA	CCCACATCCT	TCAGTCTTAA	AGTGCCTCTA	ACCACTACTT	TCCCTACCAC	1680
GTCTCCCTTT	CCAAATTCCT	CAAAAGCAAA	GCAAAACAGC	ATTCTCTGCA	CCTGTGTGCA	1740
TCGTAGCAG	TACGTAGGAG	ATCCAATATG	CCTGGCACA	GGGGTGGGG	TGGCCACAAA	1800
CAGAAGCAGC	CCTGTCTGTC	CAAGTACACG	CCTGCTGCC	ACGCACAATG	GGAGACATTC	1860
CGCAAGTTC	ACGTGATGGC	TCAGAAGAGG	GGCCTGTGAG	GAAGATGTAG	GGGCCAGCAG	1920
CCCCCGGCGC	CGCCCGCAAA	GCTGTGCTGAC	AGACGCCAGC	AGCTGCGCGG	GGCTCCGGGC	1980
GTCTCTGTCT	CCCAGGATGT	GTATCTGACT	GGAGTTTCTG	GATTAAAGGC	CAGTCTGTGC	2040
TTCATTCAC	ATCCCTGGGT	GCCCTCGGC	TCCTCTAG			

Seq ID NO: 227 Protein sequence:

Protein Accession #: XP_064321.1

65
70
75

1	11	21	31	41	51	
MYASSDQDRA	PYLPTLDKIM	PGPRLRSAR	PKAAQBPFI	EPQTYREGGG	AIVLTYALGI	60
GVGITKNTVQ	QPPQLTDSAS	IRQEDAFDNK	IDIAEDGGQT	PYBATLQGSF	QYSPTTDLFP	120
LTNGVLPFIS	MYEIQTKYQS	HNQYPNGNSK	QKTTLMRSKP	PPSTATTSTP	QTVIPKKS	180
PEVKLIKTKT	IQNGRELKFS	SLCGDLLNEV	QASEHTKSKH	BSRKEKRRKP	KXHSSTSSSE	240
RKSHKIPKLE	PREQNRPFNR	VHTISEKPRE	DPVLKEEAPV	QPLSSVPTT	EVSTGVKPVQ	300
GDLVMSKVTV	TPCMVPRLEG	RRSHHCSSCL	EILVLVPALS	LKRSFVMSL	KPLTSTGKQK	360
PTFKQTAQMG	WSPMASTTNV	SLLLGHWBGT	DQMSRGPBEP	GGRFVWVHQ	KPQIRISICH	420
RPGBPLRLS	FLRCVERRI	SSLATSQCGW	CSPPDHVCEK	CLSDYAGRRH	LTLRAQEAFL	480
GDPSRTGSLR	AVGKRYCRNS	QHRYLLQGL	LGGFLEERNA	NEYDCKLETR	EAASSTPRIP	540
YGFTHLQSE	SAPNHFFPYH	VSLSKPLKRR	ANSHPLHLCA	VVAVRRRNNM	POTRGWGHK	600
YGQPCPAKYT	PACHAQWETP	RKPHVMAQKR	GLSGRCRGOQ	PPAAPRKVAD	RRQQLPGAPG	660
CSCSQDVYLT	GVSLKASRG	FIPHPWVPPG	SS			

Seq ID NO: 228 <u>DNA sequence</u>	
Nucleic Acid Accession #: NM_006033	
5	Coding sequence: 253-1752 (underlined sequences correspond to start and stop codons)
	<div> <div>1</div> <div>11</div> <div>21</div> <div>31</div> <div>41</div> <div>51</div> </div>
10	AGCAGCGAGT CCTTGCCTCC CGGCGGCTCA GGACGAGGGC AGATCTCGTT CTGGGCAAG 60 CGTTTGACAC TCGCTCCCTG CCACGCGCCG GCGTCCGTGC CGCCAGTTT TCATTTTCCA 120 CCTTCTCTGC TCACAGTCCC CCAGCCCTTG GCGGAGAGAA GGGTCTTACC GCGCGGAT 180 GCTGGAAACA CCAAGAGGTG GTTTTGTGTT TTTAAACTT CTGTTTCTTG GAGGGGGTG 240 TGGCGGGGCA GGAAGAGCAA CTCCGTTCCT CTGCTCTGTT TCTGGAGCCT CTGCTATTGC 300 TTGCTTGGGG GAGGCTCCGT ACCTTTTGGT CCAGAGGGAC GGTGGGAAGA TAAGCTCCAC 360 15 AAACCCAAAG CTACACAGAC TGAGGTCAAA CCATCTGTGA GGTTTAACCT CCGCACCTCC 420 AAGGACCCAG AGCATGAAGG ATGCTACCTC TCCGTGGGCC ACAGCCAGCC CTTAGAAGAC 480 TGCACTTTCA ACATGACAGC TAAACCTTT TTCACTATT ACAGATGGAC GATGAGCGGT 540 ATCTTTGAAA ACTGGCTGCA CAAACTCGTG TCAGCCCTGC ACACAAGAGA GAAAGACGCC 600 AATGTAGTGT TGGTTGACTG GCTCCCTCTG GCCCACCAGC TTACACCGGA TGGCGTCAAT 660 20 AATACCAAGG TGGTGGGACA CAGCATTGCC AGGATGCTCG ACTGGCTGCA GGAGAAGGAC 720 GATTTTTCTC TGGGGAATGT CCACCTGATC GGCTACAGCC TCGGAGCGCA CGTGGCGGG 780 TATGCAGGCA ACTTGTGTGA AGGAACGGTG GCGCGAATCA CAGTTTGGGA TCCTGCGGG 840 CCATGTTTGA AAGGGGCGGA CATCCACAAG AGGCTCTCTC CGGACGATGC AGATTGTTG 900 25 GATGCTCTCC ACACCTACAC GCGTTCCTTC GGCTTGAGCA TTGTTATTCA GATGCTGTG 960 GGCACATTTG ACATCTACCC CAATGGGGGT GACTTCCAGC CAGGCTGTGG ACTCAAGGAT 1020 GTCTTGGGAT CAATTGCATA TGGAAACAAT ACAGAGGTGG TAAATGTGA CATGAGGGA 1080 GCGCTCCACC TCCTTGTGTA CTCTCTGGTG AATCAGGACA AGCGAGTTT TGCTTCCAG 1140 TGCACCTGAT CCAATCGCTT CAAAGGGGG ATCTGTCTGA GCTGCCGCAA GAACGCTGT 1200 AATAGCATTT GCTACAATGC CAAGAAAATG AGGAACAAGA GGAACAGCAA AATGTACCTA 1260 30 AAAACCCGCG CAGGCACTGC TTTCAGAGTT TACCATTATC AGATGAAAT CCATGCTTTC 1320 AGTTACAGAA ACATGGGAGA AATTGAGCCC ACCTTTTACG TCACCTTTTA TGGCACTAAT 1380 GCAGATTCCC AGACTCTGCC ACTGGAAATA GTGGAGCGGA TCGAGCAGAA TGCCACCAC 1440 ACCTTCTCTG TCTACACCGA GGAGGACTTG GGAGACCTCT TGAAGATCCA GCTCACTTGG 1500 35 GAGGGGGCCT CTCACTCTTG GTACAACTTG TGGAGGAGT TTGTCAGCTA CCTGTCTCAA 1560 CCCCCTCAAC CCGGACGGGA GCTGAATATC AGGCGCATCC GGGTGAAGTC TGGGGAACCC 1620 CAGCGGAAGC TGACATTTTG TACAGAGGAC CCTGAGAAC ACGAGATATC CCGAGCGCG 1680 GAGCTCTGTT TTGCAAGTGT TCGGATGGC TGGAGGATGA AAAAAAGAAC CAGTCCCAT 1740 GTGGAGCTTC CCTTGGGGTG CCGGGGCAAG TCTTGCCAGC AAGGACAGCA GACTTCTGTC 1800 40 TATCCAAGCC CATTGAGGAA AGTTACTGCT GAGGACCCAC CCAATGGGAG GATTCTCTC 1860 AGCCTTGACC CTGGAGCACT GGGACAACCT GGTCTCTGTT GATGGCTGGG ACTCCTCGCG 1920 GAGGGGGGAT GCGCTGCTAT AGCTCTTGCT GCTCTCTTGG AATAGCTCTA ACTCCAAACC 1980 TCTGTCCACA CCTCAGAGC ACCAAGTCCA GATTGTGTG TAAGCAGCTG GGTGCTGGG 2040 GCGCTCTGTT CACACTGGAT TGGTTCTTCA GTTGCTGGGC GAGCCTGTAC TCTGCTGAC 2100 45 GAGGAAGGCT GGTCTCGAAG AGGCGCTGTG TAGAAGGCTG TCAGCTGCTC AGCCTGCTTT 2160 GAGCCTCACT GAGAAGTCTT TCCGACAGGA GCTGACTCAT GTCAGGATGG CAGGCTGGT 2220 ATCTGTCTGG GCGCTAGCT GTTGGGGTTC TCATGGGTTG CACTGACCAT ACTGCTTACG 2280 TCTTAGCCAT TCCGCTCTGC TCCCCAGCTC ACTCTCTGAA GCACACATCA TTGGCTTTCC 2340 TATTTTCTTG TTCATTTTTT AATTGAGCAA ATGTCTATTG AACACTTAAA ATTAATTAGA 2400 50 ATGTGGTAAT GGACATATTA CTGAGCCTCT CCAATTGGAA CCCAGTGGAG TTGGGATTTT 2460 TAGACCTCTT TTCTGTTTGG ATGGTGTATG TGTATATGCA TGGGGAAGAG CACTCGGGC 2520 CTGGGGGAGG CTATAGGATA TAAGCAITAG GGAACCTGAG GCTTTAAGTG GTTCTATT 2580 CTCTTATGTT ATTATGTGCC ACCTCTTAG TTTATATGTG CCACCTCCCC TATGAGTGAC 2640 GTGTTTATGC ACTAGCAGAA TAGCAAGCAG AGTATCATTG ATGCTGGGGC CAGAAATGAT 2700 55 GCGCGTTGCC AGATATAACT GCTTGGAGC AAATCTCTTC TGTTTAGAGA GATAGAAGTT 2760 ATGACATATG TAATACACAT CTGTGTACAC AGAAACGGGC ACCTGGCAGA CAGAGCTGGT 2820 TCTAAGATT TAAACAGTGC TTTTCTCTCT CTTTGAAATA TTTTACTTTA ATACCAATGC 2880 CTTTCTCTGT TGAACCTCTT GGAAGGCCA CCAATCTTAG ATCTTGATT TGAATTAATC 2940 ACACAAATAT TGAGACACTT ACACCTTTCA AAAGATTGTT GTATGCAATT CCAATTTAGA 3000 60 GTAGGGGGAG AAGGCAACT ATTATTATCC CTATTTTACA AAATGAGGC TTAGTGAGGT 3060 TCAGCCACAT GCTAGACTT ATATACTAGT TAGTGGTGCA GCCAGGGAGA GGACTCAGAT 3120 TTCTGGAGG CAAAGTCTAT CTCTGAAACT CCAATGAGAC TTTTGCAGCC AGTTCCACC 3180 AATATGCCCC AGACGTGAGA CAAACAAGGA CTTTTTTTT TATATAGAGC CATCCATAA 3240 ATCTTAGGCC CTTTATTAA TGTATAACCA GGAGAACATC TGTGCCAAGC GTTGGACTTT 3300 TTTAGCTGAG GATTGGGGAG GAAGTGTGAC ACCAAGCAGG AGAGGAAGAA TGATTCTCTT 3360 65 TGTACTTAGG TTTTCTAAGG ACATGTTTTT AATCTGTATC GTGCCAAAGT TGTATCACTG 3420 TTAACCTACT GAAGACATAA CCAATTGAGT CTATTTTCAA GATATGTTCT CAAGCCAATT 3480 GTGTGCTTCT CTGTTTCTG TGAATGCTTT CTAGCCAAAG CGAAGCTTGT ACAGGTTGAG 3540 TATCCCTTAT CCAAAATGCT TGGAAACAGA AGTGTTCATA ATTTTAGATT ATTTTCAGAT 3600 70 TTTGGAATGT TTGCATATC ATATGTAGAT ATTTTGGGAA TAGGACCGGA GCCTAAACAC 3660 AAAATCTCAT GATGTGTGAG TTACACCTTA TCCACATAGC CTGAGGTTAA TTTTATACGA 3720 TATTTTAAAT AGTTGTGTAC ATGAAGCATG GTTTGTGTA ACTATGTGA GGGGTTTCC 3780 CATTTTGTGT CTGTTGGTG TCACAAAAGT TTTGGAATTT GGAGCAATTC GGATTTTGA 3840 TTTTGGATT AGGGTTGCTC AACCCATATT ATGGGCTGTA CATCTGGTC ACTTCTGACT 3900 75 TCTGTTTTTA CTAATGGAAG CTTTGCA
Seq ID NO: 229 <u>Protein sequence:</u>	

Protein Accession #: NP_006024.1

```
1      11      21      31      41      51
|      |      |      |      |      |
5 MSNSVPLLCF NSLCYCFAAG SPVPFGPEGR LEDKLKPEA TQTEVKPSVR FNLRTSKDPE 60
HEGCVLSVGH SQPLEDCSPN MTAKTPPIIH GNTMSGIFEN WLHKLVSAH TREKDANVVV 120
VDWLPLAQQL YTDVNMTRV VGHSTARMLD WLQEKDDPSL GSVHLIGYSL GAHVAGYAGN 180
FVKGTVGRIT GLDPAGPMFE GADIKRLSP DDADFVDVLH TYTRSPGLSI GIQMPVGHID 240
IYPNGGDFQP GCGLNDVLGS IAYGITEVV KCEHERAVHL FVDSLVAQDK PSFAPQCTDS 300
10 NRPKKGICLS CRKRCNSIG YNAKQMRNKR NSQMYLAKTR GMPTFVYHYQ MKIHVFSYKN 360
NGEIEPTFTV TLYGTNADSQ TLPLEIVERI EQMATNTPLV YTEEDLGDLL KIQITWEGAS 420
QSWYNLWKEF RSYLSQPRNF GREINIRIR VRSGETQRL TPCTEDPENT SISPGRELMF 480
RKCRDGRMK NETSPTVELP
```

15

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent
5 application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

- 1 1. A method of detecting an angiogenesis-associated transcript in a cell in
2 a patient, the method comprising contacting a biological sample from the patient with a
3 polynucleotide that selectively hybridized to a sequence at least 80% identical to a sequence
4 as shown in Tables 1-8.
- 1 2. The method of claim 1, wherein the biological sample is a tissue
2 sample.
- 1 3. The method of claim 1, wherein the biological sample comprises
2 isolated nucleic acids.
- 1 4. The method of claim 3, wherein the nucleic acids are mRNA.
- 1 5. The method of claim 3, further comprising the step of amplifying
2 nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1 6. The method of claim 1, wherein the polynucleotide comprises a
2 sequence as shown in Tables 1-8 .
- 1 7. The method of claim 1, wherein the polynucleotide is labeled.
- 1 8. The method of claim 7, wherein the label is a fluorescent label.
- 1 9. The method of claim 1, wherein the polynucleotide is immobilized on
2 a solid surface.
- 1 10. The method of claim 1, wherein the patient is undergoing a therapeutic
2 regimen to treat a disease associated with angiogenesis.
- 1 11. The method of claim 1, wherein the patient is suspected of having
2 cancer.
- 1 12. An isolated nucleic acid molecule consisting of a polynucleotide
2 sequence as shown in Tables 1-8.
- 1 13. The nucleic acid molecule of claim 12, which is labeled.
- 1 14. The nucleic acid of claim 13, wherein the label is a fluorescent label

- 1 15. An expression vector comprising the nucleic acid of claim 12.
- 1 16. A host cell comprising the expression vector of claim 15.
- 1 17. An isolated polypeptide which is encoded by a nucleic acid molecule
2 having polynucleotide sequence as shown in Tables 1-8
- 1 18. An antibody that specifically binds a polypeptide of claim 17.
- 1 19. The antibody of claim 18, further conjugated or fused to an effector
2 component.
- 1 20. The antibody of claim 19, wherein the effector component is a
2 fluorescent label.
- 1 21. The antibody of claim 19, wherein the effector component is a
2 radioisotope.
- 1 22. The antibody of claim 19, which is an antibody fragment.
- 1 23. The antibody of claim 19, which is a humanized antibody
- 1 24. A method of detecting a cell undergoing angiogenesis in a biological
2 sample from a patient, the method comprising contacting the biological sample with an
3 antibody of claim 18.
- 1 25. The method of claim 24, wherein the antibody is further conjugated or
2 fused to an effector component.
- 1 26. The method of claim 25, wherein the effector component is a
2 fluorescent label.
- 1 27. The method of detecting antibodies specific to angiogenesis in a
2 patient, the method comprising contacting a biological sample from the patient with a
3 polypeptide which is encoded by a nucleotide sequence of Tables 1-8.